Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 1-14

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.8756.1067

Bridging the Gap: A Comparative Analysis of Iranian EFL and ESP Teachers' Beliefs and Characteristics

Hassan Soodmand Afshar[1*]
Farzaneh Kasraee[2]
Maryam Jalilikia[3]

[*1]Professor of Applied Linguistics, Department of English Language and Literature, Faculty of Persian Literature and Foreign Languages, Allameh Tabataba'i University, Tehran, Iran

[2]M.A. in Applied Linguistics, Department of English Language, Bu-Ali Sina University, Hamedan, Iran

[3]Ph.D. in Applied Linguistics, Department of English Language and Literature, Ferdowsi University of Mashhad, Mashhad, Iran

ARTICLE HISTORY

Received: 16 October 2024
Revised: 26 November 2024
Accepted: 10 December 2024
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: h.soodmand@atu.ac.ir

ABSTRACT

The present study attempted to comparatively investigate English as a Foreign Language (EFL) teachers' and English for Specific Purposes (ESP) teachers' beliefs about language learning and teaching and their distinctive characteristics. To this end, 150 EFL and ESP teachers, selected randomly from various universities in Iran, were asked to fill out two pilot-tested validated five-point Likert-scale questionnaires, one on *beliefs* (including 45 items) and the other on *characteristics* (comprising 16 items). To do multiple-level analysis and triangulate the data for validation purposes, 25 teachers, selected randomly from among the study participants, were also interviewed. The results of (Quantitative data analysis such as Independent Samples t-tests, Chi-Square and Pearson product-moment correlation analyses, as well as qualitative content analysis of interviews) found no significant mismatch between ESP teachers and EFL teachers' beliefs overall; however, significant differences were found between characteristics of ESP teachers and EFL teachers overall as well as between the *specific* beliefs and also between *specific* characteristics of the two groups of the teachers. The results are discussed in detail in the paper, and the implications, which are significant for the field of language education and ESP, and the pertinent implications are presented.

KEYWORDS: Teacher beliefs; Teacher characteristics; ESP teachers; EFL teachers; Language teaching

1. Introduction

Teacher beliefs are widely recognized as a crucial factor influencing instructional decision-making, classroom practices, and student learning outcomes. In the field of English language teaching, teachers' beliefs shape their approaches to pedagogy, curriculum design, and assessment. While extensive research has explored the beliefs of English as a Foreign Language (EFL) teachers (e.g., Borg, 2006a; Phipps & Borg, 2009; Soodmand Afshar & Ghasemi, 2017), less attention has been devoted to English for Specific Purposes (ESP) teachers. Given the distinct nature of these two instructional contexts—EFL focusing on general language acquisition and ESP addressing domain-specific language needs—it is essential to investigate whether their pedagogical beliefs differ and how these differences might impact teaching practices.

Despite the growing body of research on teacher cognition, comparative studies examining the beliefs of EFL and ESP teachers remain limited (Rajabi et al., 2011; Watson, 2003). Most studies in ESP primarily focus on curriculum design, material selection, and learner needs, neglecting the role of teacher beliefs in shaping instructional decisions. This gap in the literature

Cite this article: Soodmand Afshar, H. , Kasraee Nejad, F. and JaliliKia, M. (2025). Bridging the gap: A comparative analysis of Iranian EFL and ESP teachers' beliefs and characteristics. *Applied Linguistics Inquiry*, *3*(1), 1-14. doi: 10.22077/ali.2025.8756.1067

underscores the need for a comparative analysis to determine whether and how EFL and ESP teachers differ in their beliefs about language teaching and learning.

The present study aims to address this gap by examining and comparing the beliefs and characteristics of Iranian EFL and ESP teachers. Specifically, it investigates the most and least frequently reported beliefs and characteristics, explores potential mismatches between the two groups, and discusses the implications of these differences for teacher education and professional development. The findings of this study can contribute to a deeper understanding of teacher cognition in language education and inform the design of teacher training programs tailored to the specific needs of EFL and ESP instructors.

2. Literature review

Understanding teacher beliefs and professional characteristics is essential in language education, as these factors shape instructional practices, decision-making, and student learning outcomes. While extensive research has explored teacher cognition in general language teaching (Borg, 2006a; Phipps & Borg, 2009; Soodmand Afshar & Ghasemi, 2017), comparatively little is known about how English as a Foreign Language (EFL) and English for Specific Purposes (ESP) teachers differ in their beliefs and professional characteristics. This section reviews relevant literature on teacher beliefs, theoretical perspectives, and the distinctive characteristics of EFL and ESP teachers, highlighting the research gap this study aims to address.

2.1. Teacher beliefs: definition and importance

Teacher beliefs refer to implicit or explicit assumptions that guide teachers' perceptions, decision-making, and instructional behaviors (Kagan, 1992; Lan & Lam, 2020; Pajares, 1992). According to Puchta (1999), beliefs serve as guiding principles that influence individuals' actions, shaping their responses to educational challenges and their teaching methodologies. These beliefs influence classroom management, pedagogical strategies, and interactions with students (Williams & Burden, 1997). Bodur (2003) further defines teacher beliefs as personally held opinions regarding the nature of teaching, learning, students, subject matter, and the overall classroom context. These deeply ingrained perceptions shape how teachers design their lessons, interact with learners, and implement instructional strategies.

In the context of language teaching, teacher beliefs shape how instructors approach lesson planning, error correction, assessment, and student engagement (Farrell, 2019). For instance, beliefs about language learning difficulty can influence teachers' expectations and attitudes toward students (Horwitz, 1985). Similarly, research suggests that reflective practice can help teachers critically evaluate and modify their beliefs to enhance their teaching effectiveness (Basturkmen et al., 2004; Ramezani, 2014; Rashidi & Moghadam, 2014).

2.2. Theoretical perspectives on teacher beliefs

Several theoretical frameworks explain the formation and evolution of teacher beliefs (Abelson, 1979; Farrell, 2019; Pajares, 1992).:

- Cognitivist Perspective: Beliefs are viewed as stable mental constructs that shape decision-making and are not easily altered (Li, 2013).
- Interactionist Perspective: Beliefs evolve through experience, social interactions, and professional engagement with students and colleagues (Skott, 2001).
- Theories-in-Action Perspective: This framework distinguishes between espoused beliefs (what teachers claim to believe) and beliefs-in-use (actual classroom practices), suggesting that teachers may act differently from their stated beliefs (Argyris et al., 1985; Rahimi & Sahragard, 2019).

These perspectives highlight the fact that beliefs are dynamic and context-sensitive, making it essential to study them within specific teaching environments.

2.3. Beliefs of EFL and ESP teachers

Previous research has extensively examined the beliefs of EFL teachers, particularly regarding communicative language teaching (CLT), grammar instruction, and error correction (Basturkmen et al., 2004; Lee, 2009). EFL teachers typically emphasize fluency, interaction, and general language proficiency (Borg, 2006a; Fang, 1996).

In contrast, ESP teaching is tailored to specific professional or academic needs, requiring instructors to balance linguistic competence with subject-specific knowledge (Clark & Peterson, 1986; Rajabi et al., 2011). Research suggests that ESP teachers often prioritize content knowledge over language proficiency, leading to potential differences in pedagogical beliefs (Doğruer et al. 2010; Erkmen, 2012; Watson, 2003). However, studies directly comparing the beliefs of EFL and ESP teachers are scarce, making this an important area for further investigation.

2.4. Distinctive characteristics of language teachers

In addition to beliefs, teachers' professional characteristics, such as subject knowledge, pedagogical skills, and instructional strategies, play a critical role in shaping their effectiveness. Borg (2006a) identifies five key characteristics that distinguish language teachers from other educators:

- Nature of subject matter: Language is both the content and the medium of instruction.
- Interaction patterns: Language teaching requires dynamic engagement and real-time communication (Borg, 2006a).
- Continuous knowledge development: Teachers must stay updated on linguistic, pedagogical, and technological advancements (Borg, 2006a).
- Professional isolation: Many language teachers work independently, limiting interdisciplinary collaboration (Borg, 2006a).
- Reliance on external resources: Authentic materials, technological tools, and supplementary resources are essential in language teaching (Borg, 2006a).

Given these factors, the professional identity and instructional practices of EFL and ESP teachers may vary significantly, reinforcing the need for comparative research.

3. Significance of the study and research questions

Teachers play a pivotal role in shaping students' learning experiences and outcomes. As Galluzzo (2005) states, "nothing is more central to students' learning than the quality of the teacher" (p. 142). Understanding teachers' beliefs and characteristics is essential for enhancing instructional effectiveness, designing evidence-based professional development programs, and informing language education policies (Brown, 2000; Wright et al., 1997). Research suggests that teachers' beliefs influence not only their instructional choices but also their attitudes toward student learning and professional development (Pajares, 1992; Puchta, 1999; Williams & Burden, 1997). Additionally, Al-Osaimi and Wedell (2014) argue that teacher beliefs are shaped by contextual variables, including cultural norms, institutional expectations, and students' learning goals, all of which influence pedagogical decision-making. Moreover, Riley (1997) highlights those beliefs affect key psychological factors such as motivation, attitudes, and learning strategies, ultimately shaping students' engagement and success in the learning process (Allebone & Davies, 2000).

Despite a growing body of research on teacher cognition, limited attention has been paid to the comparative study of EFL and ESP teachers' beliefs and characteristics. While EFL instruction emphasizes general language proficiency, ESP focuses on discipline-specific language use, requiring instructors to balance linguistic competence with subject-specific knowledge (Rajabi et al., 2011; Watson, 2003). However, there is a lack of empirical evidence regarding whether and how these two groups of teachers differ in their pedagogical beliefs and professional attributes. Theories of teacher cognition suggest that beliefs are shaped by professional training, teaching experiences, and contextual factors, making comparative studies particularly valuable (Borg, 2006a; Farrell, 2019).

This study aims to bridge this gap by systematically examining the beliefs and characteristics of Iranian EFL and ESP teachers. Specifically, it investigates whether significant differences exist in their reported beliefs about language teaching and learning, as well as their self-perceived professional characteristics. Additionally, it explores the extent to which these differences (if any) impact their instructional practices. By addressing these questions, the study contributes to the broader field of teacher cognition and provides insights into the professional development needs of EFL and ESP instructors. The findings will be valuable for policymakers, curriculum designers, and teacher educators seeking to enhance the quality of language instruction across diverse educational contexts.

To achieve these objectives, the study is guided by the following research questions:

1. What are the most and least frequently reported pedagogical beliefs among ESP and EFL teachers?
2. What are the most and least frequently reported professional characteristics of ESP and EFL teachers?
3. To what extent do ESP and EFL teachers differ in their overall pedagogical beliefs?
4. To what extent do ESP and EFL teachers differ in specific pedagogical beliefs, as measured by individual items in the belief questionnaire?
5. To what extent do ESP and EFL teachers differ in their overall professional characteristics?
6. To what extent do ESP and EFL teachers differ in specific professional characteristics, as measured by individual items in the characteristics' questionnaire?

4. Methodology

4.1. Participants

The present study participants included 225 ESP and EFL teachers, out of whom 150 teachers (70 ESP teachers and 80 EFL teachers) completed and returned the questionnaires. Out of this total number (150), 71 were female and 79 were male. The teachers in both groups had 1 to 30 years of teaching experience, with the highest number falling within the 1-5 year-experience category for EFL teachers (n=27) and the 6–10-year category for ESP teachers (n=20). Most teachers (n=57) were 31 to 40 years old, and most (n=87) held a doctoral degree. The participants were randomly selected from various universities throughout the country to better represent the population of teachers.

4.2. Instruments

The present study used a mixed method of data collection and analysis. That is, it adopted two types of instruments: Two Likert-scale structured questionnaires and a semi-structured interview, which are described below:

4.2.1. Likert-scale structured questionnaires

Two separate five-point Likert-scale questionnaires were adopted for data collection purposes. One questionnaire focused on EFL/ESP teachers' beliefs (BQ), and the other concentrated on ESP/EFL teachers' distinctive characteristics (CQ). The response options ranged from 1 to 5, showing strongly disagree to strongly agree.

The first part of the BQ sought some background information about teachers' gender, age, university degree, teaching experience, etc. The key items of the BQ were chosen from different accredited questionnaires in the field, including the Beliefs About Language Learning Inventory (BALLI) developed by Horwitz (1987). Because BALLI was used initially with ESL students, a modified version of it, adapted and applied to foreign language teachers by Vibulphol (2004), was used in this study. It comprises five factors in foreign language acquisition: aptitude, difficulty learning, nature of language learning, learning and communication strategies, and motivation. Another questionnaire from which some of the items of the BQ of the study were adapted was the one devised by Fives and Buehl (2005), which includes teachers' beliefs about teaching ability and pedagogical knowledge. The CQ used in the present study is provided by Borg (2006a) with only minor modifications. The questionnaires were validated through a pilot study, expert judgment, and factor analysis. BQ and CQ can be found in Appendix A and Appendix B.

4.2.2. Pilot study

In the process of the pilot study, the questionnaires were distributed among a similar group of 65 teachers at different universities (except those used in the study). Most of the answers were completed and collected via electronic mail; some were in paper and pencil format. After collecting the first round of data, the reliability and construct validity of both questionnaires were estimated. Based on the ideas and comments of the pilot study participants and the views of two experts in the field, some changes were made to the questionnaires to make them more reliable and valid. In the pilot process, the questionnaire on beliefs contained 50 items, among which five ambiguous or ill-functioning items were deleted. Also, the questionnaire used by Borg (2006a) had 18 items originally, two of which were removed for the same reason mentioned above, leaving it with 16 questions.

The Cronbach's alpha internal consistency reliability of both questionnaires was calculated. According to the reliability statistics, BQ and CQ yielded a Cronbach's alpha of 0.89 and 0.75, respectively, which revealed that they both enjoyed acceptable internal consistency reliability indices. Also, the KMO sampling adequacy measures (0.73 and 0.71 for BQ and CQ, respectively) were adequate. Additionally, the results of Bartlett's Test of Sphericity were significant (=000) for both questionnaires, which, together with the results of KMO mentioned above, confirmed the validity of the questionnaires.

4.2.3. Interview

A semi-structured interview was also designed to obtain detailed information on the questionnaire data for multiple-level analyses and triangulation purposes. Drawing on Sawani's (2000) and Barcelos' (2000) interview questions regarding teacher beliefs and characteristics and taking the views of two experts in the field in this regard into consideration, the interview was constructed and conducted with 12 ESP and 13 EFL teachers selected randomly out of the population of the study whose informed consents were also obtained. The interviews were audio recorded (with teachers' permission), transcribed, and content analyzed. As a result, the common themes and recurring patterns were extracted, coded, quantitized, and frequency analyzed. The interview questions are in Appendix C.

4.3. Procedure

First, the study participants were informed about the nature and purpose of the study and asked to complete the BQ and CQ. In the second step, twenty-five teachers selected, as mentioned above, were interviewed to examine the data obtained from the

questionnaires further. Finally, the data collected from the questionnaires and interviews were analyzed to provide the answers to the research questions.

5. Results

5.1. Results of the questionnaires

The first research question investigated the most and least frequently reported beliefs by ESP teachers and EFL teachers. The results are provided in Tables 1 and 2, respectively.

Table 1. The most frequently reported beliefs by both ESP and EFL teachers

Rank	Number of items in the questionnaire	Mean	SD
1	28	4.54	.72
2	27	4.52	.79
3	39	4.50	.73
4	40	4.48	.84
5	41	4.46	.65

Table 2. The least frequently reported beliefs by both ESP and EFL teachers

Rank	Number of items in the questionnaire	Mean	SD
1	19	2.07	.93
2	4	2.24	.98
3	17	2.42	1.05
4	10	2.43	1.10
5	5	2.76	.94

The second research question explored the most and least frequently reported characteristics of ESP and EFL teachers, the results of which are presented in Tables 3 and 4, respectively.

Table 3. The most frequently reported characteristics by both ESP and EFL teachers

Rank	Number of items in the questionnaire	Mean	SD
1	1	3.80	.81
2	14	3.78	.80
3	9	3.70	.85
4	5	3.56	.82
5	3	3.53	.96

Table 4. The least frequently reported characteristics by both ESP and EFL teachers

Rank	Number of items in the questionnaire	Mean	SD
1	12	2.68	1.08
2	13	3.06	.96
3	11	3.10	1.05
4	8	3.22	.97
5	2	3.28	.98

The third research question was whether a significant mismatch existed between ESP teachers and EFL teachers' beliefs. To answer this question, an Independent Samples *t*-test was conducted. However, the descriptive statistics for both groups of teachers are presented in Table 5 first.

Table 5. Descriptive statistics for ESP teachers and EFL teachers' beliefs

	N	Mean	Std. Deviation
EFL teachers	80	167.75	14.54
ESP teachers	70	171.62	14.01

As mentioned, an independent samples t-test was run to compare ESP and EFL teachers' beliefs. The results are presented in Table 6.

Table 6. Independent Samples t-test comparing ESP teachers' and EFL teachers' beliefs overall

	Levene's Test for Equality of Variances		t-test for Equality of Means				
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Equal variances assumed	.099	.75	-1.65	148	.100	-3.87	2.34
Equal variances not assumed			-1.66	146.61	.099	-3.87	2.33

As the results in Table 6 show, *overall*, there was no significant mismatch between ESP teachers and EFL teachers concerning their beliefs, t (148) = -1.65, p = 0.1> 0.05 (two-tailed).

However, one-way Chi-Square analyses were conducted to investigate whether there were any significant differences between ESP teachers' and EFL teachers' *specific* beliefs as measured by individual questionnaire items. Only the results of those significantly different items are presented in Table 7 to save space and for brevity.

Table 7. Chi-Square analyses comparing ESP teachers and EFL teachers on their specific beliefs

N0.	Percent		Pearson Chi-Square	df	Sig. (2-sided)
	EFL teachers	ESP teachers			
1	37.5 Agree	50 Agree	20.13	4	.000
3	60 Agree	34.3 Agree	11.13	4	.02
4	40 Disagree	44.3 Neither agree nor disagree	10.77	4	.02
17	45 Disagree	35.7 Disagree	13.75	4	.008
22	43.8 Agree	72.9 Strongly agree	22.77	4	.000
23	57.5 Agree	35.7 Agree	13.97	4	.007
25	36.3 Agree	51.4 Agree	8.38	3	.03

As the results in Table 7 show, there were significant differences between the two groups on seven individual items of BQ; that is, items number 1 (p = .000), 3 (p =.02), 4(p = .02), 17 (p =.008), 22 (p =.000), 23 (p =.007), and 25, p = .03 < 0 .05 (two-tailed).

The fifth research question of the study aimed to investigate whether there was any significant mismatch between ESP teachers and EFL teachers' overall characteristics. An Independent Samples *t*-test was conducted, the results of which are summarized in Tables 8 and 9 below:

Table 8. Descriptive statistics for ESP teachers and EFL teachers' characteristics overall

	N	Mean	Std. Deviation
EFL teachers	80	55.46	6.79
ESP teachers	70	52.58	6.38

Table 9. Independent Samples t-test comparing ESP teachers and EFL teachers' characteristics overall

	Levene's Test for Equality of Variances		t-test for Equality of Means				
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Equal variances assumed	.11	.74	2.66	148	.009	2.87	1.08
Equal variances not assumed			2.67	147.22	.008	2.87	1.07

As the results in Table 9 indicate, there was a significant mismatch between ESP teachers' and EFL teachers' characteristics *overall*, t (148) = 2.66, p = .009 < 0.05 (two-tailed).

To answer the sixth research question as to whether there was any significant mismatch between ESP teachers and EFL teachers' *specific* characteristics as measured by individual items of the questionnaire, one-way chi-square analyses were conducted, the results of which are presented in Table 10 below:

Table 10. Chi-Square analyses comparing ESP teachers and EFL teachers on their specific characteristics

N0.	Percent		Pearson Chi-Square	df	Sig. (2-sided)
	EFL teachers	ESP teachers			
1	55 Agree	41.4 Neither agree nor disagree	9.67	4	.04
5	51.3 Agree	50 Neither agree nor disagree	12.84	4	.01
9	46.3 Agree	50 Agree	12.64	4	.01
10	38.8 Agree	55.7 Neither agree nor disagree	18.49	4	.001
12	32.5 Disagree	38.6 Neither agree nor disagree	10.98	4	.02

As shown in Table 10, there were significant differences between the two groups on five specific characteristics; that is, items number 1 (p = .04), 5 (p =.01), 9 (p = .01), 10 (p =.001), and 12, p = .02 < 0.05 (two-tailed).

5.2. Results of the semi-structured interview

Tables 11 and 12 present the results of the participants' responses to the questions posed in the semi-structured interviews regarding teacher beliefs and characteristics.

Table 11. Common patterns of participants' responses to interview questions on teacher beliefs

Number of Question	Common patterns of responses	Participants			
		EFL Teachers		ESP Teachers	
		Frequency	Percent	Frequency	Percent
1.	Yes	10	76.92	7	58.34
	No	3	23.08	5	41.66
2.	Yes	6	46.16	4	33.33
	No	7	53.84	8	66.67
3.	Yes	7	53.84	5	41.66
	No	6	46.16	7	58.34
4.	Depends	2	15.39	4	33.33
	No	11	84.61	8	66.67
5. A. The most manageable parts of learning English	Grammar	1	12.5	3	23.07
	Reading	8	50	5	46.15
	Vocabulary	2	18.75	2	15.39
	Speaking	2	18.75	2	15.39
B. The most difficult parts of learning English	Writing	5	41.17	6	50
	Pronunciation	3	23.52	3	21.43
	Listening	2	11.77	3	21.43
	Speaking	2	11.77	1	7.14
	Grammar	2	11.77	0	0
6.	Depends on the goal	4	31.25	1	7.69
	All	4	31.25	4	38.46
	Vocabulary	3	19.25	1	7.69
	Grammar	2	18.25	0	0
	Reading	0	0	2	15.39
	Writing	0	0	1	7.69
	Listening	0	0	2	15.39
	Speaking	0	0	1	7.69
7.	Interest	2	13.63	2	17.39
	Prestige	1	10.63	1	10
	Academic needs	4	30.28	4	34.78
	Job	3	22.73	3	21.73
	Communication	3	22.73	2	16.1
8.	It is better	7	53.85	8	66.66
	Not necessarily	6	46.15	4	33.34
9.	They are effective	11	84.61	7	58.34
	Not much important	2	15.39	5	41.66
10.	It can be learnt	6	46.15	2	16.66
	It is innate	3	23.08	6	50
	Both of them	4	30.77	4	33.34
11.	Subject knowledge	5	25.8	5	35.71
	Teaching knowledge	4	23	4	28.57
	Psychology knowledge	3	19.35	3	21.42
	Knowledge about students	3	19.35	2	14.3
	Classroom management knowledge	2	12.9	0	0

Table 12. Common patterns and recurring themes of participants' responses to Interview questions on teacher characteristics

		Participants			
		EFL Teachers		ESP Teachers	
No. of the Question	Common patterns of responses	Frequency	Percent	Frequency	Percent
1.	Using the same medium (e.g., English) makes it easier	7	50	5	41.66
	Using the same medium makes it harder	2	14.28	1	8.34
	Language learning is a skill	2	14.28	1	8.34
	Other fields are more formal and serious	3	21.44	2	16.66
	Language is easily forgotten	0	0	1	8.34
	The amount of information is higher in language learning	0	0	2	16.66
2.	Helpful	4	30.77	4	33.34
	Effective	4	30.77	5	41.66
	More successful	5	38.46	3	25
3.	Higher	3	23.07	2	16.67
	The same	6	46.16	8	66.66
	Lower	4	30.77	2	16.67
4.	Yes	10	76.92	5	41.66
	No idea	3	23.08	7	58.34
5.	A subject in English	5	38.46	10	83.34
	English with specific content	8	61.54	2	16.66
6.	EFL teachers	4	30.76	1	8.33
	ESP teachers	1	7.7	5	41.68
	EFL teachers are familiar with specific content	4	30.76	2	16.66
	ESP teachers are good at English	2	15.39	3	25
	Both of them can teach it	2	15.39	1	8.33

6. Discussion

The first and second research questions investigated the most and least frequently reported beliefs and characteristics of ESP and EFL teachers. As the results indicated in Tables 1 to 4, some of the beliefs and characteristics were commonly shared by the teachers in both groups.

Both groups had similar beliefs regarding teachers' common beliefs about general pedagogical and pedagogical content knowledge. That is, they believed studies of teachers' classroom behavior combined with studies of pedagogical knowledge of teachers would help significantly in understanding the teaching process, which is in keeping with the general belief that with professional growth and practice, there will be changes in teachers' cognition towards more effective teaching (Bullough, 1991; Calderhead, 1991; Clift, 1991). In the first group, which included beliefs about foreign language aptitude, the majority of the participants in both groups neither agreed nor disagreed that "women are better than men at learning foreign languages", which was one of the least frequently reported beliefs of teachers.

As an example of one of the most frequently reported characteristics by teachers, it was revealed that 55% of EFL teachers and 54% of ESP teachers agreed with the first item of the questionnaire (i.e., errors being seen as a natural and desirable learning process). Although the difference is insignificant, it might imply that EFL teachers are more familiar with the role of errors in language teaching than their ESP counterparts. Borg (2006a) holds that one of the differences between language teaching and other fields is that in subjects such as mathematics and science, learners learn and apply formulae without searching for their underlying rationale; however, in language teaching, most of the learners ask their teachers to explain the rationale behind, for instance, grammatical rules (i.e., item 11 of the CQ). As an example of the least frequently reported characteristics by teachers in the present study, only 40% of EFL teachers and 37% of ESP teachers agreed with this stance. Some teachers stated that every field of study had its own methods and difficulties.

The third research question investigated any significant mismatch between ESP and EFL teachers' beliefs. As the results showed, overall, there was no such mismatch between the beliefs of these two groups of teachers. The second group of items in BQ discussed beliefs about the difficulty of language learning. The purpose was to know whether and to what extent teachers believed in the difficulty of learning language skills (i.e., listening, speaking, reading, and writing) and sub-skills or components (i.e., grammar, vocabulary, and pronunciation).

Most teachers in each group (about 84%) agreed that some languages were easier to learn than others. Similarly, most pre-service teachers in Vibulphol's (2004) study agreed with this idea; however, this percentage was much higher in the present study. Regarding the difficulty of English, more than half of the participants in each group agreed that it is a language of medium difficulty. One of the ESP teachers stated that the difficulty of teaching a foreign language "depends on the teacher, his/her knowledge and ability to teach." The Thai pre-service teachers in the study of Vibulphol (2004) had different ideas from those found in the present study. They mentioned the alphabet system was one reason they saw English as a complex language.

According to Horwitz's (1985) study, when teachers underestimate the difficulty of the language they are teaching, they might develop unrealistic expectations and regard their students as less successful. Regarding the participants' ideas about which skills and components are easy or difficult to learn, nearly half of them agreed that 'reading' was the easiest skill to acquire. 'Writing' was mentioned as the most challenging skill by both groups. However, 'pronunciation' was believed to be the second most challenging component by EFL teachers, but 'pronunciation and listening' jointly were regarded so by ESP teachers.

The third group of the items on beliefs (i.e., beliefs about the nature of language learning) investigated what was important in learning a foreign language. About half of the participants in both groups believed it necessary to know about English-speaking cultures in order to speak English. The majority of the teachers in both groups (nearly 80%) agreed that "It is best to learn English in an English-speaking country" However, as the reports of the interviews showed, although not most, some of the EFL teachers believed that "English should not necessarily be learnt in an 'all-English' environment (e.g., in UK, USA, Canada, etc.)". In contrast, ESP teachers perceived the environment as being of crucial importance.

Although learning the vocabulary of another language was naturally expected to be significantly more important for ESP teachers, on the whole, only 47% of them agreed with this belief. This percent for EFL teachers was slightly lower (i.e., 43.8%). With the belief that "the most important part of learning a foreign language is learning the grammar", 43.8% of EFL teachers and only 31.4% of ESP teachers disagreed. As Horwitz (1985, 1987) believes, having a (strong) belief in the role of vocabulary and grammar in learning English will harm the language learning process. Similarly, Peacock's (2001) study also found that less proficient EFL learners believed in learning only vocabulary and grammar, and more proficient ones disagreed with this belief.

In the group of beliefs about learning and communication strategies, about half of the participants in both groups believed that speaking English with excellent pronunciation was important. Also, nearly 80% of teachers in each group stated that they enjoyed practicing English with native speakers. More than 70% of teachers in both groups appreciated guessing as a communication strategy, and the majority disapproved of the statement, "You should not say anything in English until you can say it correctly."

Unsurprisingly, most of the participants agreed with all statements in the group of beliefs about motivations and expectations, which illustrates that learning English was important and that they tended to learn it for various reasons, especially for instrumental inspirations. Weinstein and Mayer (1986) state that the goal of learning strategies is to "affect the learner's motivational or affective state or how the learner selects, acquires, organizes, or integrates new knowledge" (p. 315). Also, according to the experiments reported by Vernon (1971) about motivation, motivation might make perception and learning easier, rapid and exact and can thus affect or be affected by one's belief system.

Concerning beliefs on the reasons for learning English, the majority responded that it was for such utilitarian purposes as getting a good job, accessing information from around the world, entering a higher education level, and helping communicate with people from other countries. In general, the results of the BQ, up to here, converge with those of similar studies, especially that of Büyükyazi (2011), who investigated the beliefs about language learning of 156 EFL students and 19 EFL teachers in Turkey. Also, with some slight variations in the results of some items, the general outcome of this study is in line with the results gained from the teachers' beliefs' study conducted by Ghobadi Mohebi and Khodaday (2011), who compared Iranian university students' and teachers' beliefs about language learning.

The rest of the questions focus on teachers' pedagogical knowledge. With the belief that individuals are born with the ability to teach, most participants neither agreed nor disagreed, and some disagreed. Most teachers believed that teaching required both innate talent and pedagogical preparation and that individuals had to develop their natural abilities to be teachers. This finding corroborates the results of Fives and Buehl (2005) in this respect.

As the results previously showed, the responses of EFL and ESP teachers to the BQ were not significantly different, which contrasts the findings of the study by Rajabi et al. (2011) conducted in the context of the present study. This might be because the beliefs of language teachers are complicated (Farrell & Ives, 2014), and a multitude of factors might be involved in the issue by different individuals, even in the same cultural and educational context.

The fourth research question dealt with whether any significant mismatch existed between ESP teachers' and EFL teachers' specific beliefs. As the results showed, significant differences were found between the two groups on seven specific beliefs measured by individual items of the questionnaire; that is, items number 1, 3, 4, 17, 22, 23, and 25. Items 1, 3, and 4 belong to the group of 'beliefs about foreign language aptitude'. This group aimed to investigate whether (EFL and ESP) teachers believed foreign language learning required specific abilities and whether some learners could learn a foreign language (i.e. English) better than others. Although Horwitz (1985, 1987) argues that these beliefs are harmful to the process of language learning for both teachers and learners, the majority of the participants in the study reported holding these kinds of beliefs. This calls for further research to see whether this has an adverse impact on teachers' classroom practices and students' achievement in various cultures and contexts.

Almost all ESP teachers in the study believed it was better to start learning English in childhood. One of the reasons for the disagreement of a small number of EFL teachers in the study with this idea might have been that they considered individual features and differences of the learners in the learning process more important than age, as one of them stated in the interview. As Puchta (1999) maintains, belief in aptitude for language learning should be given due attention. Teachers' expectations about students' learning abilities might influence their interactions with them. When a student is seen as a low-ability learner, he/she might have less interaction (both verbal and non-verbal) with the teacher in comparison with his/her competent counterparts. As

a result, these types of students might eventually not succeed, something which the teacher might also expect. This line of reasoning supports the finding of Doğruer et al. (2010), who found EFL teachers believed the most important factor affecting them was their belief in aptitude for language and how it affected their teaching styles.

Teachers did not consider translation important. Almost all ESP teachers considered repetition and practice more important in learning; however, only 71% of EFL teachers agreed. One of the possible reasons for this might be that EFL teachers regard other factors, such as watching movies and listening to music, as being more important in the process of learning a foreign language. One of the plausible reasons for EFL teachers' significantly higher beliefs in the need for speaking skills compared to their ESP counterparts might be that the latter feel English is mainly needed for reading texts, and they do not sense a need to speak the language. With item 25, which dealt with belief in and preference for native speaker friends, 36% of EFL teachers and 51% of ESP teachers agreed, which caused a significant difference. This rate was expected to be higher for the EFL group. However, ESP teachers consider having native-speaker friends more important than their EFL counterparts. The fifth research question addressed whether there existed any significant mismatch between ESP teachers' and EFL teachers' characteristics overall, the results of which indicated, overall, there was a significant difference between the two groups of teachers in this regard.

Most teachers in both groups had no idea whether the range of competing methodologies and methodological shifts in language teaching over the years outweighed similar phenomena in other subject areas. Teachers' lack of knowledge about other fields might be one of the reasons why many students are not satisfied with ESP courses. One of the differences the teachers believed existed between language teaching and other fields was that in the former, the subject and the medium for teaching are the same, a finding in line with the results of Borg (2006a). Nearly 40% of ESP teachers stated they had no idea about the many sources for language learning and teaching. The high agreement of EFL teachers (61%) with item 7, which stated, "driven by powerful commercial forces, language teaching is characterized by a proliferation of teaching and learning resources unparalleled in other subjects", caused a significant difference between the two groups. One of the possible reasons for these differences might be that ESP teachers only know the English language and are not thus familiar with language teaching theories, methodologies and practices.

With the feature that language teaching has a practical outcome that is not a characteristic of other subjects, 40% of ESP teachers had no idea, and only 45% of them agreed, while 65% of EFL teachers agreed. This might show that EFL teachers believe more in utilitarian purposes and instrumental reasons for teaching and learning .EnglishIn response to the sixth research question as to whether there was any significant mismatch between ESP teachers and EFL teachers' specific characteristics measured by individual items of the questionnaire, the results showed significant differences between the two groups on five specific characteristics; that is, items number 1, 5, 9, 10, and 12.

One of the possible reasons for the existence of a significant difference between the answers of the two groups to item 9 appears to be that some of the ESP teachers might not have been familiar with the concepts of declarative and procedural knowledge. Moving forward, with item 12 (i.e., language teachers are considered low-status compared with their subject-matter counterparts), 38% of ESP teachers and 18% of EFL teachers neither disagreed nor agreed, and 34% of ESP teachers and 32% of EFL teachers disagreed. One of the likely reasons for this disagreement with this statement might be that, in the academic situation of universities, teachers of various disciplines consider themselves at the same level. However, the results here are not in line with those of Borg (2006a), in which the EFL teachers felt isolated and low status compared to their counterparts in other fields.

The findings of this study, highlighting the significant differences in specific beliefs and characteristics between EFL teachers and ESP teachers, align with prior research emphasizing the contextual sensitivity of teacher beliefs (Argyris et al., 1985; Farrell, 2019). This disparity underscores the importance of tailored teacher training programs that address these contextual variances. For instance, the stronger emphasis on procedural knowledge among EFL teachers compared to the declarative focus observed in some ESP counterparts, Borg (2006b) suggests the necessity for differentiated professional development strategies. Moreover, the results resonate with Puchta's (1999) assertion that beliefs about learners' aptitude can shape teacher-student interactions, potentially reinforcing or mitigating students' learning anxieties. These findings also echo Peacock's (2001) conclusion that teacher beliefs about grammar and vocabulary can significantly influence language learning approaches, indicating a need for pedagogical alignment with contemporary language acquisition theories. Policymakers and curriculum designers should consider integrating reflective practices into teacher training, enabling educators to critically evaluate their beliefs and align them with effective instructional practices, as Farrell and Ives (2014) suggested. Such initiatives can bridge the gap between espoused theories and classroom realities, fostering a more effective language-teaching ecosystem.

7. Conclusion and implications of the study

This study investigated and compared the beliefs and characteristics of EFL and ESP teachers at the tertiary educational level in Iran. The results indicated that there was no significant mismatch between their beliefs. More specifically, the results indicated that most of the teachers in both groups believed in the role of aptitude in foreign language learning, considered English as a language of medium difficulty, and regarded reading as the easiest and writing as the most difficult skill to learn. Learning the culture of the foreign language and speaking in the specific environment where the given language is spoken were considered important by both ESP and EFL teachers in general. Both groups of teachers considered the practice as an important learning strategy. They believed the motivation behind learning English was mainly instrumental and included getting a good job, accessing information from around the world, entering higher education levels, etc. They also believed effective teachers should

have high pedagogical knowledge about the theoretical foundations and implications of teaching practices, the subject matter they teach, how to motivate and engage students, how to deliver information, etc., to mention only a few. However, ESP and EFL teachers had significantly different ideas about learning a foreign language's most important skills/ components.

Significant differences were also found between the two groups of teachers concerning their characteristics. The results further indicated that many ESP teachers, for instance, were unfamiliar with English language teaching. Additionally, the results showed how EFL and ESP teachers thought differently about who should teach the ESP courses and what should be taught in these courses (a subject in English or English with specific content). The results of this study emphasize the complex interplay between beliefs and characteristics in shaping the pedagogical approaches of EFL teachers and ESP teachers. While no overarching mismatch was found between the two groups' beliefs, notable differences in specific beliefs and characteristics highlight the context-dependent nature of teaching practices. These findings suggest that professional development programs should adopt a differentiated approach, addressing EFL teachers and ESP teachers' unique needs to enhance their effectiveness. Additionally, the study underscores the importance of fostering reflective practices among educators, enabling them to align their beliefs with evidence-based methodologies. Future research could further explore how these differences impact student outcomes and investigate the role of institutional and cultural factors in shaping teacher beliefs and practices.

Moreover, the present study's findings suggest some implications for teaching and learning English. Although teaching at the university level shows that teachers might have reached the professional growth needed and have a good command of the theories in their fields, it does not necessarily mean that they need no evaluation of their salient beliefs and characteristics. The evaluation of (belief systems and characteristics) can make teachers reflect on their teaching, enabling them to become aware of the differences between their current beliefs and practices and those leading to favorable learning conditions and then change their beliefs, attitudes, knowledge, and methods. As Farrell and Ives (2014, p.14) maintain, by bringing beliefs to the level of conscious awareness and articulating them, teachers can develop a kind of self-evaluation which can ultimately lead to "a form of self-mediated professional development".

The results of the evaluation of teacher beliefs can also help educational policymakers and administrators in making better decisions and teacher educators in designing more effective pre-service and in-service training courses in which the improper written-in-the-stone beliefs of teachers on teaching and learning a foreign language are challenged and those in line with optimal foreign language learning conditions are highlighted and strengthened if we are to improve the current situation.

8. References

Abelson, R. P. (1979). Differences between beliefs and knowledge systems. *Cognitive Science, 3,* 355–366. https://doi.org/10.1207/s15516709cog0304_4

Al-Osaimi, S. & Wedell, M. (2014). Beliefs about second language learning: the influence of learning context and learning purpose. *The Language Learning Journal, 42*(1), 5-24. https://doi.org/10.1080/09571736.2012.661753

Allebone, B., & Davies, D. (2000). Developing subject knowledge. In S. Herne, J. Jessel, & J. Griffiths (Eds.), *Study to teach: A guide to studying in teacher education* (pp. 60–69). Routledge.

Argyris, C., Putnam, R., & McLain, S. D. (1985). *Action science: concepts, methods, and skills for research and intervention*. Jossey-Bass.

Barcelos, A. M. (2000). Understanding teachers' and students' language learning beliefs in experience: A Deweyan approach. Unpublished doctoral dissertation, College of Education, University of Alabama, Tuscaloosa.

Basturkmen, H., Loewen, S. & Ellis, R. (2004). Teachers' stated beliefs about incidental focus on form and their classroom practices. *Applied Linguistics, 25*(2), 243-272. https://doi.org/10.1093/applin/25.2.243

Bodur, Y. (2003). Pre-service teachers' learning of multiculturalism in a teacher education program. *Electronic Theses, Treatises and Dissertations, 3635*.

Borg, S. (2006). *Teacher Cognition and Language Education*. Continuum.

Borg, S. (2006). The distinctive characteristics of foreign language teachers. *Language Teaching Research, 10*(1), 3–31. https://doi.org/10.1191/1362168806lr182oa

Brown, H. D. (2000). *Principles of language learning and teaching* (4th ed.). Addison Wesley Longman, Inc.

Bullough, R. V. (1991). Exploring personal teaching metaphors in pre-service teacher education. *Journal of Teacher Education*, *42*(1), 43–51. https://doi.org/10.1177/002248719104200107

Büyükyazi, M. (2011). The beliefs of university foreign language students and their teachers about language learning. *Sosyal Bilimler, 8*(1), 169-182.

Calderhead, J. (1991). The nature and growth of knowledge in student teaching. *Teaching and Teacher Education, 7*(5), 531–535. https://doi.org/10.1016/0742-051X(91)90047-S

Clark, C. M., & Peterson, P. L. (1986). Teachers' thought processes. In M. Wittrock (Ed.), Handbook of research in teaching (3rd ed., pp. 255–296). Macmillan.

Clift, R. T. (1991). Learning to teach English—maybe: A study of knowledge development. *Journal of Teacher Education*, *42*(5), 357–372. https://doi.org/10.1177/002248719104200505

Doğruer, N., Meneviş, I., & Eyyam, R. (2010). Telling ELT Tales out of School, EFL Teachers' beliefs on learning English and their teaching styles. *Procedia Social and Behavioral Sciences, 3*, 83–87. https://doi.org/10.1016/j.sbspro.2010.07.015

Erkmen, B. (2012). Ways to uncover teachers' beliefs. *Procedia Social and Behavioral Sciences, 47*, 141 – 146. https://doi.org/10.1016/j.sbspro.2012.06.628

Fang, Z. H. (1996). A review of research on teacher beliefs and practices. *Educational Research*, *38*(1), 47–65. https://doi.org/10.1080/0013188960380104

Farrell, T. (2019). Standing on the Shoulders of Giants: Interpreting Reflective Practice in TESOL. *Iranian Journal of Language Teaching Research, 7*(3), 1–14. https://doi.org/10.30466/ijltr.2019.120733

Farrell, T. S., & Ives, J. (2014). Exploring teacher beliefs and classroom practices through reflective practice: A case study. *Language Teaching Research*, published online July 25 2014. https://doi.org/10.1177/1362168814541722

Fives, H. & Buehl, M. M. (2005). Assessing teachers' beliefs about pedagogical knowledge: Developing an instrument. Paper presented at the Southwest Educational Research Association Annual Meeting, New Orleans, LA.

Galluzzo, G. (2005). Performance assessment and reviewing teacher education: The possibilities of NBPTS standards. *Academic Research Library, 74*, 142–145. https://doi.org/10.3200/TCHS.78.4.142-145

GhobadiMohebi, S., & Khodaday, E. (2011). A study on Iranian university students' and teachers' beliefs about language learning. *Iranian EFL Journal, 7*, 98-110.

Horwitz, E. K. (1985). Using students' beliefs about language learning and teaching in the foreign language methods class. *Foreign Language Annals, 18*(4), 333–340. https://doi.org/10.1111/j.1944-9720.1985.tb01811.x

Horwitz, E. K. (1987). Surveying student beliefs about language learning. In A. Wenden & J. Rubin (Eds.), *Learner strategies in language learning* (pp. 119- 129). Prentice-Hall International.

Hos, R. & Kekec, M. (2014). The Mismatch between Non-native English as a Foreign Language (EFL) Teachers' Grammar Beliefs and Classroom Practices. *Journal of Language Teaching and Research*, *5*, 80-87. http://dx.doi.org/10.4304/jltr.5.1.80-87

Kagan, D. M. (1992). Professional growth among pre-service and beginning teachers. *Review of Educational Research, 64*, 129–169. https://doi.org/10.3102/00346543062002129

Lan, W. & Lam, R. (2020). Exploring an EFL Teacher's Beliefs and Practices in Teaching Topical Debates in Mainland China, *Iranian Journal of Language Teaching Research*, *8*(1), 25-44. https://doi.org/10.30466/ijltr.2020.120806

Li, L. (2013). The complexity of language teachers' beliefs and practice: one EFL teacher's theories. *The Language Learning Journal, 41*(2), 175–191. https://doi.org/10.1080/09571736.2013.790132

Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal,* 63(1), 13–22. https://doi.org/10.1093/elt/ccn010

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307-332. https://doi.org/10.3102/00346543062003307

Peacock, M. (2001). Pre-service teachers' beliefs about second language learning: A longitudinal study. *System, 29*, 177–195. https://doi.org/10.1016/S0346-251X(01)00010-0

Phipps, S. & Borg, S. (2009). Exploring tensions between teachers' grammar teaching beliefs and practices, *System, 37*(3), 380-390. https://doi.org/10.1016/j.system.2009.03.002

Puchta, H. (1999). Beyond materials, techniques and linguistic analyses: The role of motivation, beliefs and identity. *Plenary session at the LATEFL: 33rd International Annual Conference, Edinburgh*, 64–72.

Rajabi, P., Kiany, Gh. R., & Maftoon, P. (2011). Iranian English major vs. subject-matter ESP teachers' beliefs and instructional practices in ESP classes: A comparative study. 1st International Conference on Foreign Language Teaching and Applied Linguistics, 1017–1028. Retrieved April 25, 2013, from http://eprints.ibu.edu.ba/573/1/FLTAL%202011%20Proceed%C4%B1ngs%20Book.pdf

Rahimi, S. A. & Sahragard, R. (2019). The Mismatch between EAP Teachers' Beliefs and Classroom Practices toward Formative Assessment in the EFL Context of Hormozgan, *American Journal of Humanities and Social Sciences Research*, *3*(1), 40-49.

Ramazani, M. (2014). Mismatches in Beliefs between Teachers and Students, and Characteristics of Effective English Teacher: An Iranian Context, *Social and Behavioral Sciences*, *98*, 1518–1527. https://doi.org/10.1016/j.sbspro.2014.03.573

Rashidi, N. & Moghadam, M. (2014). The Discrepancy between Teachers' Belief and Practice, from the Sociocultural Perspective. *Studies in English Language Teaching*, *3*(3), 2015. http://dx.doi.org/10.22158/selt.v3n3p252

Riley, P. (1997). The guru and the conjurer: Aspects of counselling for self-access. In P. Benson & P. Voller (Eds.), *autonomy and Independence in Language Learning* (pp.114–31). Longman.

Richards, J. C., & Lockhart, C. (1994). *Reflective teaching in second language classrooms*. Cambridge University Press.

Sawani, F. M. (2000). Beliefs about language learning and strategy use: A case study of a Libyan learner. Retrieved February 15, 2013, from webspace.webring.com/people.

Skott, J. (2001). The emerging practices of a novice teacher: the roles of his school mathematics image. *Journal of Mathematics Teacher Education*,*4*(1), 3–28. https://doi.org/10.1023/A:1009978831627

Soodmand Afshar, H. & Ghasemi. Z. (2017). EFL Teachers' Beliefs and their Actual Classroom Practices: Any Difference? *Iranian Journal of Applied Language Studies*, *9*(2), 173–208. https://doi.org/10.22111/ijals.2017.3547

Vernon, M. D. (1971). *The psychology of perception*. Penguin Books Ltd., Baltimore, MD.

Vibulphol, J. (2004). Beliefs about language learning and teaching approaches of pre-service EFL teachers in Thailand. Unpublished doctoral dissertation, University of Oklahoma.

Watson T. R. (2003). EAP or TEAP? *Journal of English for Academic Purposes*, *2* (2), 147–156.

Weinstein, C. & Mayer, R. (1986). The teaching of learning strategies. In M.C. Wittrock (Ed.), *Handbook of Research on Teaching* (pp. 315-327). Macmillan.

Williams, M. & Burden, R. (1997). *Psychology for Language Teachers: A Social Constructivist Approach.* Cambridge University Press.

Wright, S. P., Hom, S. P., & Sanders, W. L. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of Personal Evaluation and Education*, *11*, 57–67. https://doi.org/10.1023/A:1007999204543

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 15-30

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.8775.1072

The Role of Chat-GPT-Driven Materials in Shaping EFL Education: A Comparative Study across Iranian Language Institutes, Public Schools, and Private Schools

Moahdeseh Sadeghi[1]
Ali Mohammad Fazilatfar[2*]

[1]Ph.D. Candidate in Applied Linguistics, English Language and Literature Department, Language and Literature Faculty, Yazd University, Yazd, Iran

[*2]Professor of Applied Linguistics, English Language and Literature Department, Language and Literature Faculty, Yazd University, Yazd, Iran

ARTICLE HISTORY

Received: 18 November 2024
Revised: 25 December 2024
Accepted: 06 January 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: afazilatfar@yazd.ac.ir

ABSTRACT

Artificial Intelligence (AI) is affecting various aspects of education and there is a need to consider the consequences of this technological improvement on achievement and attitudes of the most direct stakeholders. Therefore, this study has focused on the effect of AI on teaching English with three key areas of concern: the effect of Chat-GPT-driven materials on learning outcomes, the attitudes of teachers who use Chat-GPT -informed pedagogy, and the students' attitudes towards Chat-GPT-integrated materials. This study used a quasi-experimental intervention-based research design to achieve the intended objectives. Based on the type of educational institution, the participants were divided into three groups (90 students from Iranian language institutes, 100 students from public high schools, and 100 students from private high schools). Furthermore, fourteen teachers from these institutions shared their viewpoints on the use of AI, particularly Chat-GPT, in developing instructional materials. Pre- and post-tests were administered to check the effect of treatments on each group (based on the syllabus developed for those specific educational institutions). Moreover, two researcher-developed questionnaires were administered among teachers (composed of 27 Likert-Scale items) and students (composed of 23 Likert-Scale items). The obtained results indicated that Chat-GPT significantly improved EFL learners' performance in language institutions. Teachers looked upon AI with favor. This was based on its potential to change the produced educational materials and improve students' engagement in the learning process. Students, in all three environments, looked positively toward Chat-GPT-driven materials. In short, this study showed the potential for AI to upgrade EFL teaching in Iran.

KEYWORDS: Artificial intelligence; Language institutes; Material development; Private high schools; Public high schools

1. Introduction

Nowadays, technology has significantly changed education. Technology advancements can remove the limitations imposed by geographical distance, and open up a big ocean of knowledge that people may access online (Xiao & Zhi, 2023). Amongst these Technological advances, Artificial Intelligence (AI) stands out. With its potential learning capabilities, natural language processing, and data analysis power, AI is evolving today's personalized learning systems, marking crucial innovations in this regard (Azadnia, 2024; Shaikh et al, 2023). Though the meta- frameworks defined in the algorithms of AI-driven platforms, these platforms analyze how individuals learn. By doing so, they can make various educational content (Marr & Ward, 2019; Marr, 2022; Ulla et al., 2023).

Cite this article: Sedghi, M. and Fazilatfar, A. M. (2025). The role of AI-driven materials in shaping EFL education: A comparative study across iranian language institutes, public schools, and private schools. *Applied Linguistics Inquiry*, *3*(1), 15-30. doi: 10.22077/ali.2025.8775.1072

The prompts provided to AI can help customization of the learning experience. These prompts make the ground ready for AI, and through these prompts, AI can know the tastes and proficiency levels of individual learners, in detail. Educational materials can, therefore, be fine-tuned by those needs (Nguyen, 2023). These technical innovations can facilitate education, in general, and language acquisition, in particular. These facilities enable learners to have immediate and real-time feedback, resulting in the provision of an interactive teaching tool at any moment of their life (Yan, 2023). Moreover, materials used to teach any new language, such as English, provide the fundamental principles that help the learners through the interaction. They guide learners in varied aspects of language learning such as grammar, lexis, and culture. According to Ellis (2004), materials are the signposts every participant must follow. Well-made educational materials help not only create a good linguistic foundation but also improve the students' ability to think critically, understand culture and society, and communicate effectively. Good instructional materials bridge the gap between abstract knowledge and real-life application, thus enabling learners to confidently learn the language in its various complicated contexts. Therefore, the development of AI- driven EFL materials that can support the learners during various aspects of education may ensure practicability and efficiency (Azadnia, 2024).

In spite of the potential benefits of AI in education, its integration into EFL context can be challenging (Liu & Ma, 2023). The Iranian educational system has undergone significant changes during the last decades, trying to emphasize on English proficiency, and various global studies have revealed that AI-driven materials can help the language learning stakeholders, in this regard (Marr, 2022). However, the development of AI-driven content and language learning tasks, in Iran, may encounter significant restrictions due to the limited access to digital infrastructure, financial problems, international sanctions, and low levels of technological literacy among educators (Abbasi et al., 2021).

Moreover, few studies have been carried out considering the role of new technology advancements on EFL material development in the educational context of Iran. Besides, empirical data on the influence of AI-driven materials on different ELT educational contexts, such as language institutes, public schools, and private schools, are few and rare. Furthermore, both teachers' and students' perceptions about AI-driven instructional materials remain underexplored. Therefore, there is a need to investigate the effect of AI on ELT material development and the attitude of EFL learners toward the integration of AI technology into the process of educational material development.

2. Literature review

As mentioned previously, AI has provided new ways and techniques for teaching and learning in language education through the introduction of several new capabilities that are gradually revolutionizing the field (Mitra & Banerjee, 2022). The technology of AI is yet evolving, but its applications in oral or written proficiency, motivation of learning, and assessment have obtained substantial attention. This section serves to contextualize current research into the broader discourse of AI in language education, identifying key contributions and gaps, implications for teaching practices, with a particular focus on how AI reshapes EFL learning in Iranian educational contexts.

One central theme in AI-based research on language learning is the role of AI in enhancing writing or speech recognition and its evaluation. Ali's (2020) research, therefore, focused on a five-year content analysis to establish that AI technology has significantly enhanced the accuracy of speech recognition systems, thereby further advancing the potential for more effective and efficient language assessment practices. Moreover, Ali's study highlighted that the integration of AI into approaches such as flipped classrooms contributed to higher levels of self-efficacy, improved listening skills, and increased motivation among learners. On the other hand, regarding writing proficiency, Azadnia (2024), in a review paper, focused on 28 research works concerned with the application of different forms of AI applications (specifically ChatGPT) in writing. Finally, the reviewer concluded that the body of research on writing is rich and advocates the positive effect of AI on writing enhancement among EFL learners.

In addition to the studies focused on a specific aspect of language learning process, there are some research works focused on the process of language learning, in general. In one of such studies, Leunard et al. (2023) investigated the impact of the employment of AI, ChatGPT, in a gamification approach to the process of language learning. The finding underpinned the ability of AI in creating an interactive learning environment, improving the general language proficiency of the students. Moreover, the study also revealed that while the integration of ChatGPT improved students' reading abilities, there was a tremendous increase in motivation and completion of assignments. This study was, however, limited to Arabic and gamified learning alone. So, the question is raised whether similar beneficial effects can be visible under less interactive and more linguistically diverse learning situations. Another critical perspective is given by Moulieswaran and Prasantha Kumar (2023), who researched how ESL learners feel and what kind of challenges they face in their learning processes on their way to language acquisition with AI support. While the participants generally expressed a preference for using AI tools, the study also identified several challenges, such as technological constraints and the performance of AI-based mobile applications. Wang et al. (2023) go even deeper into how AI can become an integral part of the process of language learning. In their research, they recorded interactions with AI in elementary students over a period of three months. Thereby, their research made possible the identification of different learner profiles based on student engagement with AI. What it ultimately proved was that some learners benefited from personalized guidance in AI, although others could not use it effectively. This variation thus underlines the fact that the use of AI in language education is nuanced, considering the differences between individual learners, and it needs to be complementary, not replacing, to foster a more inclusive community of learning.

Despite this now growing body of research on AI in language education, a number of gaps continue to exist. Recent research has focused on the direct application of AI tools in educational contexts (Leunard et al., 2023; Moulieswaran & Prasantha Kumar, 2023); however, due to the novelty of this technology, lack of familiarity with the device may reduce the optimized use and

achievement. Moreover, educational and assessment materials, as the grounding blocks of any learning process, are not vastly taken into account. Besides that, in most of the conducted research works, the most direct stakeholders of the educational system (i.e. teachers and students) were defectively analyzed (either concerned with students or teachers).

Given these gaps, therefore, the present study was conducted to explore in much depth the role of AI-driven instructional material in achievement of EFL learners within Iranian educational contexts. Chat-GPT was taken into account as the sample AI platform, due to its popularity and accessibility. The study tried to achieve a better understanding of how AI-driven materials may result in an insight into teachers' and learners' attitudes towards language education. This can thus go a long way toward addressing the literature gaps that currently exist and furthering our knowledge of the potential use of AI in language education.

Therefore, the present study attempted to answer the following research questions:

1. Does the application of Chat-GPT-driven materials improve the performances of EFL learners in Iranian language institutes, public high schools, and private high schools?

2. What are the attitudes of EFL teachers regarding the integration of Chat-GPT in material development for English language education in Iranian language institutes, public high schools, and private high schools?

3. What are the attitudes of EFL students regarding the integration of Chat-GPT in material development for English language education in Iranian language institutes, public high schools, and private high schools?

3. Methodology

3.1. Research design

A quasi-experimental intervention-based research design was adopted in this study in order to explore the effect of AI-driven instructional materials in EFL context under diverse Iranian educational settings. The application of AI-driven instructional materials in the form of instructional intervention among the designated groups was considered as the independent variable. Besides, the ultimate achievement of EFL learners in different educational contexts, teachers' attitudes, and students' attitudes were taken into account as the dependent variables.

3.2. Participants

The current study was conducted collecting the intended data from among the language institutes, public schools, and private schools in Yazd. The research was an attempt to realize how the participants (290 EFL learners) reacted to the intended treatments. The selected students were divided into three groups as determined by the type of school they attended, with 90 people being from Iranian language institutes, 100 learners selected from public high schools, and a further 100 EFL students chosen from private high schools. The rationale for selecting intact classes was that public high schools in Iran primarily follow textbooks, regulations, and methodologies mandated by the Ministry of Education. On the other hand, the ones in private high schools take advantage of complementary instructional materials (such as books, additional classes, videos, and audios) and a more liberal instructional syllabus (due to the less strict monitoring by the ministry). Finally, the language institutes are performing in a liberal context in which the instructional materials and teaching strategies are selected based on the viewpoints of the authorities in those language institutes.

Each classroom, at the language institutes, contained 12 to 17 students. Of these, three classes were selected as the experimental group and the remaining three were taken into account as a control group. It is important to note that three language institutes (owned by the same person but in different regions of the city) were selected for this study, with each of them contributing two classes, one assigned to the control group and the other chosen as an experimental one. The classes were randomly selected to be experimental and control groups to ensure unbiased distribution. The students, in these classes, were considered to be at or around intermediate level of English proficiency based on their performance in placement tests administered by the institutes to classify the learners in the intended classes.

Similarly, 10th-grade classes (each consisting of 20 to 25 students) of both public and private high schools were selected to participate in the study, due to their age proximity to the ones in the language institutes. In each of the schools, two classes were randomly selected as experimental groups and the other two were assigned as the control ones. The use of random selection techniques was based on the idea to minimize potential selection bias and increase the internal validity of the study. To ensure the homogeneity of the participants of both kinds of schools with the ones in language institutes (to be at or around intermediate proficiency level), their English proficiency was determined based on their most recent English exam scores and teacher evaluations. Of course, there were some outliers in each class, who were tested (since it was not possible to omit them from the intact classes), but their scores were not taken into account in this study.

Moreover, the participants were surveyed regarding their linguistic background such as their L1, prior studies in English, and residence in English speaking countries. Through the investigation, the homogeneity of the participants was proved. The learners' English learning period had ranged from three months to one year by the time of testing. Almost none of them had been living in an English-speaking community, and English could be considered as a foreign language for them.

In addition, this research covered the viewpoints of educators, which included language teachers at Iranian Language Institutes as well as the ones in public and private schools. The teacher cohort consisted of 6 instructors from language institutes, 4 teachers from public schools, and 4 teachers from private schools, all of whom contributed their insights and perceptions regarding the utilization of AI in material development. It is worth mentioning that the selected teachers were the ones involved in the instructional process during the interventions of the study.

3.3. Instruments and materials

3.3.1. AI-based applications

This study was mainly based on the instructional materials developed by Chat-GPT platform. It is worth mentioning that the major instructional materials i.e. course textbooks were the ordinary ones proposed by the ministry of education or the institute; however, the complementary materials that helped the students and teachers through the process of instruction were designed to be Chat-GPT-driven. Therefore, in order to enable the participants to develop the intended materials, there was a need to have some pre-treatment instructional period for both students and teachers. The instructional period lasted for one and half an hour (students and teachers were taught in separate sessions to avoid any discomfort in the case of problem or question).

The researcher introduced and explained about different varieties of AI applications or platforms that could help the students or teachers in addition to the major focus on the potentials of Chat-GPT platform. Teachers were provided with some information on how to craft AI-generated reading materials, exercises, and test questions. The students were taught about using AI for self-practice, self-correction, and personalized learning. In spite of the training, some challenges were observed. A few of the teachers, at the beginning, were doubtful with AI and its accuracy, reliability, validity of content, and if it could meet the curriculum's purposes. Then there were other challenges, where a number of the students found difficulty in formulating clear prompts. Therefore, sample prompts provided and more structured guidelines were added in another training session. In fact, the students and teachers were provided with the prompts that could help receive the most optimized version of the materials. For instance, the teachers were provided with the following prompt for developing a text based on the vocabulary of each chapter in the textbook.

> "write a reading text, composed of 250 to 300 words, for EFL learners at ---------- proficiency level, focused on ---------- topic, including this list of vocabulary---------- ."

On the other hand, the students were provided with the prompts related to self-practice or self-correction. The following prompt was developed as an example of the prompts for checking students' writing.

> "Please identify and pinpoint any grammatical, structural, or discoursal problem in the following text: ----------"

In addition to the prompts, various examples were provided two both groups on the potentials and capabilities of the mentioned AI platforms. It is worth mentioning that the prompts were carefully developed and validated following several steps. First, the pedagogical needs and instructional requirements of public schools, private schools, and language institutes were analyzed, in order to be able to say that the prompts addressed the curriculum goals. Then, the prompts were reviewed for effectiveness, and pedagogical suitability by language teaching experts and a specialist in computer science and AI. Finally, the prompts were tested on a small group of students and teachers. Feedback was collected to revise and optimize them.

In these three settings, the AI-generated materials were systematically integrated into classroom instruction. In public and private schools Chat-GPT-generated reading passages and exercises were developed by teachers twice a week, to supplement textbook activities. Students used the Chat-GPT for their writing assignments or self-practice outside class. In the language institutes, Chat-GPT was used for text creation in every session for vocabulary, reading, and writing improvement. Moreover, in all groups, students and teachers used AI-generated content for exam preparation, vocabulary building, interactive exercises, etc. at any time they felt the need.

3.3.2. Pre and post-test assessments

Selection of pre-test and post-test instruments was a laborious process because the tests had to be in line with the particular educational context in which each participating institution was functioning in. For the public and private high school groups, the assessments were based on official final test sanctions proposed by the Ministry of Education. The tests included reading comprehension, vocabulary use, grammar, and writing tasks, as aligned with the ministry's English curriculum. These tests were selected by the researchers, since they normally undergo very strict validation procedures in order to ensure that the tests fit the curriculum contents and learning objectives. This was well explained by a series of tests and analyses conducted by educational experts and officials. Thus, the validity and reliability of the tests were taken for granted by the wide range of procedures that validation experts and education authorities have carried out on them.

Regarding the language institute group, the tests were institutional, specially prepared to match the syllabuses and learning goals of EFL programs at that particular institute. However, while their origins lay in the institute, they proved to have considerable validity and reliability. These assessments indeed matched not only the needs of the institute but the standards themselves. These were confirmed in expert review for their ability to measure, quite effectively, the language proficiency abilities required by these tasks or conditions. Finally, these assessments also reflected good properties of reliability, as evidenced

by a Cronbach's α of 0.84, a coefficient which shows the intrinsic confidence and certainty with which their results on academic performance are accepted as true.

3.3.3. Teacher survey

Designed by researchers, the teacher survey was an AI-assisted tool with 27 items which aimed at gathering pedagogical and practical insights from EFL teachers (See Appendix A). They had about twenty to twenty-five minutes to fill it up in order to give well thought and comprehensive answers to a combination of Likert-type scale items. The focus of this questionnaire was on various aspects such as teachers' pedagogical perspectives on AI integration in materials development, their experiences with AI-enhanced content, their perceptions on the advantages and disadvantages associated with AI integration, as well as their confidence levels when it comes to utilizing AI-driven materials for teaching purposes. To ensure its reliability, the teacher survey underwent rigorous internal consistency assessments using Cronbach's alpha ($\alpha = 0.87$). Moreover, the questionnaire was developed based on prior research and expert recommendations. Then, a pilot study was conducted with 4 teachers to revise the potential challenges of the items. Finally, content validity was thoroughly established through experts' involvement in questionnaire review process, after the revisions.

3.3.4. Learner survey

Since learners are mostly observed as passive participants whose performance is mostly studies (Azadnia, 2024), while their ideas are often neglected, the present study sought to fill this gap in the literature to some extent. With only 23 Likert-type scale items, the learner survey acted as a method created by researchers for collecting information regarding EFL learners' views, preferences, and perceived impacts of AI on material development (See Appendix B). The survey aimed at enabling the learner to see through a window of his/her mind in terms of attitudes and preferences. Students had approximately fifteen to twenty minutes to fill in this questionnaire, with the aim of obtaining meaningful answers while still capturing their attention.

The scope of this study involved different aspects including learners' views on AI-enhanced materials, their tendencies when selecting between traditional and AI-based materials and their evaluations of how AI affects their general learning experiences. This survey was found to have high reliability levels after conducting Cronbach's alpha which indicated good internal consistency ($\alpha = 0.91$). Furthermore, content validity was ensured through a long process of designing very comprehensive questionnaires and subjecting them to pilot testing and critical expert reviews.

3.4. Procedure

The research enterprise started by taking a careful selection-by consulting with the supervisors in the education office of Yazd Province about the quality of instruction and the quantity of the participants in the mentioned educational centers-of the participating institutions which accounted for three Iranian language institutes besides a wide range of public and private high schools in Yazd, Iran. The targeted EFL Learners were primarily identified and included from these educational institutions. Of all the identified population sample, 90 learners comprised those studying in Iranian language institutes-with the same instructional policies-, another 100 were public high school learners, and the remaining 100 attended high schools of private facilitation. The participants in each of these educational settings, selected through cluster sampling, that is, intact classes in each institution, were further divided into experimental and control groups.

One of the most important areas was to design and develop Chat-GPT -derived material based on the curriculum and objectives of the participating institutions. Therefore, the AI-based instructional period was administered among the teachers and students, through which the participants could have developed their own materials. The materials developed by the students themselves to facilitate the practice process were not checked by the researchers since each student may have self-practice according to his taste and preference. However, the instructional materials developed by the teachers themselves were checked in collaboration with other colleagues in order to match the needs and proficiency level of the students in addition to the educational norms and regulations. In fact, the students could themselves use or ask for AI-mediated materials whenever they felt they need it. However, in the private and public high schools, in the classes, the researchers could not take such a liberal view into account. They had to obey the rules and regulations of the schools and the ministry of education. They, therefore, allowed only the teachers to use AI-mediated materials in the classes, and students were only aware that they were instructed through new ways of teaching.

The integration of Chat-GPT in material development would allow for dynamic and adaptive learning development for the Iranian language institutes. The students would have much more personalized and responsive approaches to language learning through virtual AI tutors (which would explain each of the questions proposed by them) integrated into the institutes' learning management systems. These virtual tutors applied natural language processing algorithms and assessed the response of students to exercises and quizzes. For example, if a student struggled with verb conjugation, the Chat-GPT system would instantly provide targeted exercises and detailed explanations about verb tenses. This level of accuracy enabled learners to effectively deal with their own weaknesses. Moreover, AI algorithms carefully examined each learner's unique patterns and preferences. If a student had a preference for visual learning, the Chat-GPT system would easily provide more video-based content or interactive visual exercises. On the other hand, for those who performed well with auditory cues, the system adjusted the materials to include more audio-based activities. This flexibility ensured that students interacted with the materials in ways that matched their individual learning styles.

Besides this, AI-based materials supported the principle of learner autonomy. In this respect, students could find supplementary resources and practice tasks themselves by means of Chat-GPT-driven instruments. Such aids were always made on the fly to suit every individual learner's needs. For example, if a student had some kind of problem with phrasal verbs and he would ask some questions in this regard or check his erroneous homework in the mentioned platforms, then immediately AI generated exercises and tests that focused on helping him improve that particular skill.

Chat-GPT-embedded materials in public high schools were developed to enhance traditional classroom instruction through technology applications that effectively engaged students. Language learning tasks enriched with Chat-GPT platform became useful classroom tools. Among other things, these programs included interactive exercises that enabled students to learn languages interestingly and deeply. The key point about them was their provision of immediate responses. For example, if during oral practice one mispronounced a term, the artificial intelligence system without delay detected this error giving ways of such mistakes being addressed and thus enabling learners to better their pronunciation as they go along. Besides, the complexity of exercises and quizzes adjusted automatically as students improved in their language proficiency levels, ensuring that there was always a consistent level of challenge which would keep them from being bored or overwhelmed.

In private schools, mostly like public high schools, Chat-GPT-based contents facilitated the generation of personalized teaching materials. These instructional resources were produced faster than ever before using these tools (Chat-GPT 3.5 (free version), Gemini.google.com, poe.com, and claude.ai). For example, if a chapter in the textbook was about business, the Chat-GPT content creation tool could generate a bunch of business vocabulary flashcards and some related questions like quizzes autonomously. This saved teachers time and effort that would have been used to develop everything from scratch. Besides, Chat-GPT-generated materials in private schools were intended for active learning. For instance, interactive exercises and scenarios were created to simulate authentic language use in real life situations thus making students prepared for practical English communication scenarios.

On the other hand, students in the control groups used traditional learning materials such as course books and audio resources. After a two-month period, post-tests, teacher's surveys and learner's surveys were used to collect the desired data; this was done at the end of two months of treatment. The collected data was then analyzed by the designated statistician.

4. Results

The first research question, in this study, was concerned with the effect of application of Chat-GPT-driven materials on the performances of EFL learners in Iranian language institutes, public high schools, and private high schools. The t-test results revealed differences between the mean scores of experimental and control groups, from pretest to posttest.

The participants, in public experimental group, had a significant improvement from pretest (M = 43.80) to posttest (M = 54.82). This resulted in a notable mean difference (MD = -11.024, SD = 19.376). The associated t-value was (t = -3.817, df = 44, p < .001). On the other hand, the ones in the public control group did not manifest any significant improvement from pretest (M = 43.69) to posttest (M = 46.95), yielding a smaller mean difference (MD = -3.265, SD = 18.468). The t-value for this group was (t = -1.186, df = 44, p = .242), indicating no statistically significant improvement. According to the numerical findings, obtained through independent samples t-test, there was a significant difference between the posttest scores of experimental and control groups in public high schools; t (78) = 4.28, p=.0.000). The magnitude of the difference in the means was large (Eta squared= 0.79).

Moving to the next experimental group concerned with instructional materials in language institutes, the participants demonstrated a significant enhancement from pretest (M = 42.84), to posttest (M = 68.11). This remarkable improvement resulted in a mean difference (MD = -25.266, SD = 16.794). The corresponding t-value was (t = -10.638, df = 49, p < .001). on the other hand, in the institute control group, an increase was observed from pretest (M = 44.08) to posttest (M = 51.41); however, the mean difference (MD = -7.326, SD = 18.709) and t-value (t = -2.769, df = 49, p = .008), indicated a smaller improvement rate (although significant). According to the numerical findings represented on the performance of the institute control group, although the participants had a significant improvement from pretest to posttest, the participants in the experimental group outperformed the ones control group (MD = -16.70; t (78) = 2.68, p=.0.000). The magnitude of the difference in the means was very large (Eta squared= 0.86).

Finally, the participants in private experimental group had an improvement from pretest (M = 41.32) to posttest (M = 65.90), leading to a significant mean difference (MD = -24.580, SD = 21.559) and high t-value (t = -8.062, df = 49, p < .001). On the other hand, the private control group had an improvement from pretest (M = 43.96) to posttest (M = 51.12), leading to a smaller mean difference (MD = -7.155, SD = 21.333) and t-value (t = -2.372, df = 49, p = .022), indicating a statistically significant but smaller improvement. However, it is worth noting that the participants in the experimental group outperformed the ones control group (MD = -14.78; t (78) = 2.61, p=.0.000). Moreover, the magnitude of the difference in the means was large (Eta squared= 0.81).

Furthermore, an ANCOVA was conducted to determine the effect of different groups: public experimental (Mean = 54.82, SD = 15.190), institute experimental (Mean = 68.11, SD = 14.391), and private experimental (Mean = 65.90, SD = 17.142) on posttest scores with pretest scores as the covariate. The results are shown in Table 1.

Table 1. ANCOVA results for the posttest scores in all three groups

Source	df	F	Sig.	Partial Eta Squared
group	2	1.973	.043	.128
pretest	1	.430	.513	.003
group * pretest	2	.758	.471	.011

The ANCOVA results, represented in Table 1, showed that there was a statistically significant difference ($F = 4.238$, $p < 0.001$) between the posttest scores of the three groups. Moreover, the intercept was highly significant ($F = 155.609$, $p < 0.001$), showing the difference in posttests while controlling for the pretest scores. However, no statistical numbers, in the mentioned table, could show the magnitude of the difference between the groups. Table 2 provides a two-by-two comparison of posttest scores between the groups.

Table 2. Post-hoc results for the posttest scores in all three groups

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.[b]	95% Confidence Interval for Difference[b]	
					Lower Bound	Upper Bound
publicexp	instituteexp	-13.408*	3.230	.000	-21.235	-5.581
	privateexp	-11.091*	3.239	.002	-18.940	-3.242
instituteexp	publicexp	13.408*	3.230	.000	5.581	21.235
	privateexp	2.317	3.148	1.000	-5.311	9.945
privateexp	publicexp	11.091*	3.239	.002	3.242	18.940
	instituteexp	-2.317	3.148	1.000	-9.945	5.311

The post-hoc pairwise comparisons as shown in Table 2, indicated that there are significant differences between the public experimental group and both institute experimental group ($p < 0.001$) and the private experimental group ($p = 0.002$). Nevertheless, there was no significant difference between institute experimental and private experimental groups ($p=1.000$).

The second research question, addressed in this study, was about the attitudes of EFL teachers towards AI integration in materials development for English language education in Iranian language institutes, public high schools and private high schools. ANOVA analysis was performed to establish whether different groups such as public (Mean = 109.25, SD = 15.67), institute (Mean = 133.67, SD = 8.07), and private (Mean = 111.75, SD = 14.68), have any effect on teachers' attitude by assuming that variances are equal among these groups.

Table 3. ANOVA results for teachers' attitude in all three groups

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1852.595	2	926.298	5.963	.018
Within Groups	1708.833	11	155.348		
Total	3561.429	13			

The ANOVA results recorded a statistically significant ($F = 5.963$, $p = 0.018$) difference, demonstrating that at least one group had a significant effect on teacher attitudes. Therefore, Table 4 is provided to show the two-by-two comparisons.

Table 4. Post-hoc results for teachers' attitude in all three groups

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.	95% Confidence Interval	
					Lower Bound	Upper Bound
public	institute	-24.41667*	8.04540	.028	-46.1462	-2.6872
	private	-2.50000	8.81330	.957	-26.3035	21.3035
institute	public	24.41667*	8.04540	.028	2.6872	46.1462
	private	21.91667*	8.04540	.048	.1872	43.6462
private	public	2.50000	8.81330	.957	-21.3035	26.3035
	institute	-21.91667*	8.04540	.048	-43.6462	-.1872

Post-hoc pairwise comparisons, represented in Table 4, using Tukey's test, revealed a significant difference ($p = 0.028$) between teacher attitudes in the institute group and the ones in the public group. However, no significant difference ($p = 0.957$) was found between the public and private groups. On the other hand, significant difference ($p = 0.048$) was observed in teachers'

attitudes from the institute group to the private group.

The last research question, in this study, dealt with was about the attitudes of EFL learners towards AI integration in materials development for English language education in Iranian language institutes, public high schools and private high schools. An ANOVA analysis was conducted to examine the effect of different groups (public (Mean = 96.44, SD = 16.64), institute (Mean = 90.26, SD = 19.25), and private (Mean = 93.76, SD = 20.44)) on learners' attitudes.

Table 5. ANOVA results for learners' attitudes in all three groups

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	916.521	2	458.261	1.281	.281
Within Groups	50813.851	142	357.844		
Total	51730.372	144			

The ANOVA results represented in Table 5 revealed that there was no statistically significant difference in learners' attitudes among the three groups (F = 1.281, p = 0.281). in other words, group membership did not significantly affect learners' attitudes.

5. Discussion

The present study contributed to the existing knowledge on the role of Chat-GPT-driven instructional material in achievement of EFL learners within Iranian educational contexts. Moreover, the study tried to achieve a better understanding of how Chat-GPT-driven materials may result in an insight into teachers' and learners' attitudes towards language education.

The first research question was proposed to see whether Chat-GPT-developed materials improved participants' performance in three different educational settings. There were significant differences in mean scores between experimental and control groups across a range of educational institutions when investigating the effects of AI integration into instructional material development on EFL learners' achievement. In particular, there were significant improvements in learners' posttest scores compared to their pretest scores among experimental groups using Chat-GPT-enhanced materials. This implied that AI integration had a positive effect on performance of EFL students in Iranian language institutes, public schools, and private schools.

In comparison to peers from public and private schools, language institute teachers tended to have more positive attitudes toward the integration of AI. The results show that EFL institution teachers were more in favor of using Chat-GPT-enhanced materials and they saw potential advantages in such technology for classroom use.

Besides, this study sought to assess EFL students' attitudes on how they felt about AI usage during material production. Iranian language institutes, public schools and private schools had relatively similar views on integrating it. Consequently, these three sets of learners did not differ significantly in their opinions. This showed that EFL students in various types of Iranian educational institutions had similar opinions on using Chat-GPT-driven materials, with no significant difference in their attitudes.

In summary, these research findings suggest that integrating AI into materials considerably improves performance of EFL learners. However, teachers' attitudes varied across different types of educational institutions, whereas learners' attitudes remained constant irrespective of whether the institutions were public or private based.

The findings of this study supported the results of previous studies and also underlined the potential of AI in improving language education. For example, Ali (2020) stated that AI influences several factors that are related to language education positively, including speech recognition, flipped learning, language instruction, and assessment. Equally, Leunard et al. (2023) examined the benefits of AI, specifically GPT Chat Application, regarding enhancement in Arabic Language Learning. These were in line with findings from our study, where performance of EFL learners was enhanced a great deal with AI-enhanced learning materials. Similarly, Moulieswaran and Prasantha Kumar (2023) showed that ESL students had a preference for AI-based tools, while our conclusions have also identified their inclination for the use of artificial intelligence in teaching foreign languages. However, that does not mean Moulieswaran and Prasantha Kumar disagreed with the very notion that AI programs need revision when it comes to instructing languages; it means our study focused on the optimization of AI integration, something more important than anything that may be discussed by them. More recently, Wang et al. (2023) have explored students' interactions with these agents and found a fair amount of variety among learner profiles in the AI-supported learning environment.

This study gave new perspectives in solving various problems identified in existing literature. The Chat-GPT influence on language acquisition viewed holistically embraced all aspects such as reading, writing, speaking, and listening. In this regard, the study was different from other studies, partial to show more general perspective on how it contributed to language education. This was elaborated to encompass within its coverage its application in such an area as public schools, private schools, and language institutes, among many others, in that way making the findings more relevant for different linguistic backgrounds and educational institutions. Thus, these results became more relevant and informative regarding the applicability of Chat-GPT in different educational contexts. Furthermore, the study analyzed how teachers applied Chat-GPT in their teaching methodologies.

This was a critical aspect since it outlined the idea of the teachers themselves in applying Chat-GPT to language learning, an important aspect necessary for effective implementation. The study also acknowledged that each student uses Chat-GPT while learning the language in their own way, hence they should be taught accordingly.

6. Conclusion

This study has thus added to the increasing research into AI in language education by providing a holistic view on the role of AI-driven instructional material in achievement of EFL learners within Iranian educational contexts. The findings highlighted the positive contribution of AI in EFL learners' performance, teacher perspectives, and learner attitudes. The study also highlighted how personalized AI tutoring is crucial and how AI platforms can facilitate the language learning process.

Importantly, it was indicated that even in the most constrained environments, such as public schools, AI has a positive effect on language education. However, it had a much stronger effect in less constrained contexts, such as language institutes. The basic point arising from this is that AI is appropriate for all these relatively flexible educational contexts and allowing more flexibility in the teaching and learning environment may lead to far greater improvements. The positive contribution of AI to language education, therefore, can be exploited for a whole spectrum of educational settings, though the degree of influence may vary depending on specific constraints and adaptability of each context.

Ultimately, the present study enriched our understanding of the use of AI in language education and provided further insight into ways in which educators, researchers, and policy-makers who seek to exploit AI for language learning and teaching might most effectively avail themselves of this opportunity.

7. References

Abbasi, S., Chalak, A., & Heidari Tabrizi, H. (2021). Impact of online strategies-based instruction on Iranian advanced EFL learners' speaking scores. *Journal of Modern Research in English Language Studies*, *8*(3), 51–71. http://dx.doi.org/10.30479/jmrels.2020.11916.1478

Ali, Z. (2020). Artificial Intelligence (AI): A review of its uses in language teaching and learning. *IOP Conference Series: Materials Science and Engineering*, *769*(1). https://doi.org/10.1088/1757-899X/769/1/012043

Amirjalili, F. (2024). Exploring the efficacy of ChatGPT in personalized language learning: an intervention study in Iranian ESL classrooms. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*. https://doi.org/10.22099/tesl.2024.49390.3260

Azadnia, M. (2024). ChatGPT-assisted language learning and teaching: A scoping review of research on ChatGPT use in L2 pedagogy and education. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *43*(2), 49-85. https://doi.org/10.22099/tesl.2024.49169.3250

Ellis, R. (2008). *The study of second language acquisition* (2nd edition). Oxford University Press.

Ellis, R. (2011). Task-based language teaching: Sorting out the misunderstandings. *International Journal of Applied Linguistics, 19* (*3*), *221-246*. https://doi.org/10.1111/j.1473-4192.2009.00231.x

Healey, D. (2018). Technology enhanced learning environments. In J. I. Liontas(Ed.), *The TESOL Encyclopedia of English Language Teaching*. Hoboken Wiley. https://doi.org/10.1002/9781118784235.eelt0437

Leunard, H., Rachmawati, R., Zani, B. N., & Maharjan, K. (2023). GPT Chat: Opportunities and Challenges in the Learning Process of Arabic Language in Higher Education. *Journal International of Lingua and Technology*, *2*(1), 11–23. https://doi.org/10.55849/jiltech.v2i1.439

Li, R. (2020). Using Artificial Intelligence in Learning English as a Foreign Language: an Examination of IELTS LIULISHUO as an Online Platform. *Journal of Higher Education Research*, *1*(2), 85-89. https://doi.org/10.32629/jher.v1i2.178

Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching*, 1–14. https://doi.org/1-14.10.1080/17501229.2023.2240316

Marr. B, & Ward, M. (2019). Artificial Intelligence in Practice. UK: Print.

Marr, B. (2022, October 12). *The key definitions of Artificial Intelligence (AI) explain its importance*. Forbes. Retrieved December 8, 2022, from https://www.forbes.com/sites/bernardmarr/2018/02/14/the-key-definitions-of-artificial-intelligence-ai-that-explain-its-importance/?sh=3a2dbeb14f5d

Mitra, N., & Banerjee, A. (2022). A Study on Using AI in Promoting English Language Learning. In Emerging Technologies in Data Mining and Information Security (pp. 287–297). Lecture Notes in Networks and Systems (LNNS), Volume 490. https://doi.org/10.1007/978-981-19-4052-1_30

Moulieswaran, N., & Prasantha Kumar, N. S. (2023). Investigating ESL learners' perception and problem towards artificial intelligence (AI) -assisted English language learning and teaching. *World Journal of English Language*, *13*(5), 290-290.

Nguyen, T. T. H. (2023). EFL teachers' perspectives toward the use of ChatGPT in writing classes: A case study at Van Lang University. *International Journal of Language Instruction*, 2(3), 1–47. https://doi.org/10.54855/ijli.23231

Rombot, O., Boeriswati, E., & Atwi Suparman, M. (2021). The development of blended learning model for reading comprehension learning of BIPA students in grade V elementary school. *Elementary Education Online*, *20*(5), 1155-1164. https://doi.org/10.17051/ilkonline.2021.05.126

Shaikh, S., Yayilgan, S. Y., Klimova, B., & Pikhart, M. (2023). Assessing the usability of ChatGPT for formal English language learning. *European Journal of Investigating Health, Psychology and Education*, *13*, 1937–1960. https://doi.org/10.3390/ejihpe13090140

Ulla, M. B., Perales, W. F., & Busbus, S, O. (2023). To generate or stop generating response': Exploring EFL teachers' perspectives on ChatGPT in English language teaching in Thailand. *Learning, Research and Practice*, *9*(2), 168–182. https://doi.org/10.1080/23735082.2023.2257252

Wang, X., Liu, Q., Pang, H. J., Tan, S. C., Lei, J., Wallace, M., & Li, L. (2023). What matters in AI-supported learning: A study of human-AI interactions in language learning using cluster analysis and epistemic network analysis. *Computers & Education*, *194*, 104703-104703. https://doi.org/10.1016/j.compedu.2022.104703

Xiao, Y., & Zhi. Y. (2023). An exploratory study of EFL learners' use of ChatGPT for language learning tasks: Experience and perceptions. *Languages*, *8*(212), 1–12. https://doi.org/10.3390/languages8030212

Yan, D. (2023). Impact of ChatGPT on learners in an L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11),1–25. https://doi.org/10.1007/s10639-023-11742-4

8. Appendices

8.1. Appendix A

Teacher Survey on EFL Educators' Perspectives on AI in Materials Development

Demographic Information
Gender:
1. Male
2. Female
Age:
1. 18-24
2. 25-34
3. 35-44
4. 45-54
5. 55 and above
English Proficiency Level:
1. Intermediate
2. Advanced
3. Native-like
Years of English Teaching Experience:
1. Less than 1 year
2. 1-3 years
3. 4-6 years
4. 7 or more years

Instructions: Please indicate your level of agreement with each statement by selecting the appropriate number on the scale:

	Strongly Disagree	**Disagree**	**No Comment**	**Agree**	**Strongly Agree**
Pedagogical Perspectives on the Use of AI in Materials Development:					

1. AI can enhance the effectiveness of language teaching materials.					
2. The integration of AI in materials development aligns with my pedagogical philosophy.					
3. AI-enhanced materials cater to diverse learning styles in my classroom.					
4. I feel confident in incorporating AI-based materials into my teaching practices.					
5. AI in material development positively influences student engagement in my classes.					
6. AI can provide personalized learning experiences for my students.					
7. AI-enhanced materials facilitate more efficient learning outcomes.					
Actual Encounters with Radically Rehashed Educational Content:					
8. I have encountered radically rehashed educational content that bears little resemblance to traditional materials.					
9. The rehashed educational content has positively impacted my teaching approach.					
10. Integrating radically rehashed content has improved student understanding in my classes.					
11. The use of rehashed content challenges traditional teaching norms in a positive way.					
12. Radically rehashed content enhances my students' critical thinking skills.					
Attitudes Toward Benefits and Problems of Integrating New Ideas:					
13. I see potential benefits in integrating radically rehashed educational content.					
14. Integrating new ideas challenges me to adapt my teaching methods.					
15. Radically rehashed content contributes to a more dynamic learning environment.					
16. There are potential problems associated with integrating radically rehashed content.					
17. Addressing potential problems in integrating new ideas is essential for effective teaching.					
Self-Confidence in Applying AI-Enhanced Material in Practices:					
18. I feel confident in my ability to effectively use AI-enhanced materials in my classroom.					
19. Professional development opportunities have enhanced my confidence in utilizing AI in teaching.					
20. AI-enhanced materials can positively impact student learning outcomes.					
21. Integrating AI in teaching practices aligns with my professional goals.					
22. I actively seek out ways to enhance my skills in using AI in language education.					
Reliability and Content Validity Assessments:					
23. The survey items comprehensively assess various aspects of my pedagogical perspectives.					
24. The survey adequately covers the challenges and benefits of AI in materials development.					
25. The survey items are clear and easy to understand.					
26. The survey effectively captures the experiences of EFL educators in using AI-enhanced materials.					
27. I feel confident that the survey results will contribute valuable insights to the field of language education.					

8.2. Appendix B

Learner Survey on EFL Their Perspectives on AI in Materials Development

Demographic Information

Gender:

1. Male

2. Female

Age:

1. 18-24

2. 25-34

3. 35-44

4. 45-54

5. 55 and above

English Proficiency Level:

1. Intermediate

2. Advanced

3. Native-like

Years of English Learning Experience:

1. Less than 1 year

2. 1-3 years

3. 4-6 years

4. 7 or more years

Learners' Views on AI-Enhanced Materials

1. AI-enhanced materials positively impact my language learning experience.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

2. I find AI-enhanced materials more engaging and interactive compared to traditional materials.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

3. AI-based materials help me understand and grasp language concepts more easily.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral

- 4 - Agree
- 5 - Strongly Agree

4. The use of AI in material development enhances my motivation to learn English.
- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

5. AI-enhanced materials cater better to my individual learning needs.
- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

6. AI-based materials contribute to a more personalized learning experience.
- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

7. The incorporation of AI in materials positively influences my language proficiency.
- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

8. AI in material development has improved my overall language learning experience.
- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

9. AI-based materials make language learning more efficient for me.
- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

10. AI-based materials allow for a more interactive and dynamic learning environment.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

Tendencies when Selecting Between Traditional and AI-Based Materials:

11. Traditional materials are more effective for my language learning than AI-based materials.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

12. I prefer using traditional materials over AI-based materials for language learning.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

13. Traditional materials provide a more reliable foundation for language learning.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

14. I believe AI-based materials are more up-to-date and relevant than traditional materials.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

15. Traditional materials are more enjoyable for me compared to AI-enhanced materials.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

16. I prefer a combination of traditional and AI-based materials in my language learning.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

17. AI-based materials cater to diverse learning styles better than traditional materials.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

Evaluations of How AI Affects General Learning Experiences:

18. The time spent using AI-based materials is more productive for my learning.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

19. AI in material development helps me practice language skills in a real-world context.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

20. I feel more confident in my language abilities when using AI-enhanced materials.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

21. AI in material development positively influences my attitude towards learning English.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

22. I am open to exploring new AI-based materials for language learning.

- 1 - Strongly Disagree

- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

23. Overall, I believe AI has a significant impact on the future of language education.

- 1 - Strongly Disagree
- 2 - Disagree
- 3 - Neutral
- 4 - Agree
- 5 - Strongly Agree

Thank you for participating in this survey!

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 31-43

Applied Linguistics Inquiry

https://doi.org/10.22077/ali. 2025.9022.1108

Exploring AI-based Collaborative Reflective Practice in Light of ChatGPT: Insights from EFL Preservice Teachers

Mohammad Hossein Arefian[1]
Rajab Esfandiari[2*]
Abbas Ali Zarei[3]

[1]Ph.D. Candidate in Applied Linguistics, Department of English Language Teaching, Faculty of Literature and Humanities, Imam Khomeini International University, Qazvin, Iran

[*2,3]Professor of Applied Linguistics, Department of English Language Teaching, Faculty of Literature and Humanities, Imam Khomeini International University, Qazvin, Iran

ARTICLE HISTORY

Received: 27 November 2024
Revised: 11 December 2024
Accepted: 30 January 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: esfandiari@hum.ikiu.ac.ir

ABSTRACT

Few studies have explored how English-as-a-foreign-language (EFL) preservice teachers can do collaborative reflective practice (CRP) to enhance their knowledge and skills, using artificial intelligence (AI) tools. The main purpose of this study, adopting a basic interpretive qualitative design, is to investigate how CRP can be implemented with ChatGPT to develop preservice EFL teachers professionally. After selecting the participants through a purposive sampling procedure and exploring the perspectives of eight preservice EFL teachers and two teacher educators through narratives, interviews, observations, and group discussions, the data were analyzed through thematic analysis. The results of data analysis showed that employing CRP implemented in ChatGPT has the potential to provide several benefits, including the ability to contemplate new experiences, connect theory and practice, acquire knowledge and expertise, become critical and autonomous, embrace challenges and opportunities, improve professionally, socially, and emotionally, regulate emotions and thoughts, and foster creativity. Furthermore, the use of ChatGPT in CRP can function as a virtual mentor and collaborator to engage with preservice teachers, find problems and resolve issues, and enhance professional growth. The findings of the present study may carry implications for teacher educators and policy makers to enhance social and reflective practices as they integrate AI into teacher education programs to enhance preservice teachers' cognitive and social functioning.

KEYWORDS: Artificial intelligence; ChatGPT; Collaborative reflective practice; Pre-service teachers; Teacher education

1. Introduction

Previously, teacher education programs focused on disseminating abstract knowledge, theories, and methods through explanations and descriptions within a top-down approach (Sweeting & Carey, 2025). More recently, bottom-up teacher education programs concentrate on student teachers' inquiry-based, active, independent, technology-oriented, and collective learning (Ahn & Kwon, 2025; Farrell, 2015). Through this mode of teacher education, pre-service EFL teachers can learn how to improve their knowledge, skills, and practices continuously via reflections and collaborations actively when they become novice teachers. Therefore, reflective practice (Gudeta, 2022) can help preservice EFL teachers observe and monitor their practices systematically, learn from doing practically, foster their awareness theoretically, notice their gaps carefully, plan for actions and implementation wisely, and have greater agency personally. Accordingly, they become able to make wise decisions, transform their practices and beliefs, personalize their theories and theorize their practice, regulate their thoughts and emotions, and enhance their creativity (Farrell, 2022; Ghamoushi, 2025; Mann & Walsh, 2017; Meanwell & Kleiner, 2014). Moreover, teacher educators can encourage preservice EFL teachers to do reflective practice with other peers and colleagues collectively and collaboratively to learn how to receive constructive

Cite this article: Arefian, M. H., Esfandiari, R. and Zarei, A. A. (2025). Exploring AI-based collaborative reflective practice in light of ChatGPT: Insights from EFL preservice teachers. *Applied Linguistics Inquiry*, *3*(1), 31-43. doi: 10.22077/ali.2025.9022.1108

feedback from others, possess rich scaffolding and social support, and have more cultural and social understanding by the means of CRP (Alvarado Gutierrez et al., 2019).

CRP, conceptualized as a platform for planning, acting, observing, and reflecting collaboratively (Cole et al., 2018), can be adopted to provide foundation for integrating collaboration and interaction with individual reflective practices in a team- or peer-based professional development. Therefore, CRP can enhance reflective practices in a friendly and supportive communities of practice by observing and sharing practices, recognizing problems and finding solutions with others, boosting interactions and cooperation, and negotiating meanings, modifications, and adaptations (Cirocki & Farrell, 2017). In addition, preservice EFL teachers develop their teaching quality and experiential learning (Savoca, 2021), form a new belief system and enhance criticality (Zeichner & Liu, 2010), improve their situated and contextual understanding, and follow socio-cultural theory through CRP (Daniel et al., 2013; Modarresi & Alavi, 2014; Vygotsky, 1978). Most recently, AI tools have been regarded as a novel way of approaching and supporting education in general and teacher education in particular (Celik, 2023; Lameras & Arnab, 2021; Salas-Pilco et al., 2022). AI can include various modes of technologies for instructional purposes, such as using chabots, automated grading systems, and intelligent tutoring system (Lin et al., 2023).

For teachers, ChatGPT can be implemented as an "interlocutor, content provider, teaching assistant, and evaluator" (Jeon & Lee, 2023, p.9), leading to professional development enhancement. Likewise, ChatGPT can plan preservice EFL teachers' learning and provide them with the intended content (Kasneci et al., 2023), offer constructive feedback and practical guidance (Koedinger et al., 2012), and support their collaborative and personalized learning experiences. Furthermore, it lowers their workload by generating tasks, assessment, and lesson plans (Javaid et al., 2023), and makes a portfolio for checking their progress. Additionally, using automated scoring for their formative and summative assessment (Yuan et al., 2020), monitoring to enhance their learning and teaching quality (Celik et al., 2023), and teaching them how to use AI when they become novice teachers are other benefits.

However, ChatGPT may have drawbacks during teacher education, namely, limited social practice, analytical and reflective thinking, unethical practices, over-reliance on AI, and lack of creativity and originality (Jeon & Lee, 2023; Kasneci et al., 2023; Kohnke et al., 2023; Zeng et al., 2024). Since AI literacy is considered a crucial technological literacy that has to be developed for students and teachers at schools, preservice EFL teachers can learn how to use it in the future by implementing it with other peers and colleagues during their teacher education programs. Although AI has previously been found to be an effective tool for teachers' professional development (Kim, 2023), studies investigating how preservice EFL teachers can do CRPs with AI to enhance their knowledge and skills are very scarce, and more research studies are needed to examine AI-based CRP among preservice teachers to help better understand how it contributes to teachers' professional development. Hence, CRP can be utilized with ChatGPT to see how EFL preservice teachers can be developed and improved with this new practical technological innovation.

2. Literature review

2.1. Preservice teachers' collaborative reflective practices

Teacher education can be improved through authentic, experiential, and practical instructional experiences. It is important to involve preservice EFL teachers in reflective and collaborative practices to learn from their real instructional practices. These reflective and collaborative activities function as critical opportunities for them to strengthen their knowledge of teaching content, obtain effective experience, and reform educational practices. By nurturing reflective and collective professional development and supporting contextualized understanding, preservice teachers will be better prepared to fulfill the developing needs of their students (Borko, 2004; Guskey, 2002). As stated previously, "a reflective approach to language teaching is one in which teachers collect data about their teaching, examine their attitudes, beliefs, assumptions and teaching practices, and use the information obtained as a basis for critical reflection about their teaching" (Richards & Lockhard, 1994, p. 15). Hence, as a systematic and rigorous mode of thinking, reflective practice can help preservice teachers to enhance their awareness and intellectual development (Rodgers, 2002).

Moreover, they can experience a meaning-making process by reflecting on practices, thinking analytically, connecting theory and practice, and learning in a community of practice with peers and colleagues. In this case, reflective practice can indicate preservice EFL teachers how to move from intuitive and impulsive decision-making to a more systematic and sound approach in their actions and interpretations when they become a novice teacher at school (Farrell, 2012). Preservice EFL teachers need to be familiar with reflective practice to monitor their learning and teaching practices and evaluate their teaching to develop professionally. Consequently, they acquire skills to notice their strengths and problems, possess a deeper thinking style, question and modify their thoughts and actions appropriately, reach greater self-awareness, and generate novel solutions (Kim & Lee, 2002). These benefits of reflective practice can enhance the pre-service EFL teachers' understanding of professional development during their teaching practices in the future.

Reflective practice is an intellectual learning process that can be complemented by "a set of attitudes in which teachers systematically collect data about their practice, and while engaging in dialogue with others, using the data to make informed decisions about their practice both inside and outside the classroom" (Farrell, 2015, p.123). To prevent individualistic and isolated reflection, pre-service teachers can develop their reflective professional development through collaboration, leading to more collegial scaffolding and feedback, collective inquiry, maximized learning potential in a community, and mutual negotiation of thoughts and practices (Putman et al., 2025). Preservice EFL teachers need to practice working with other peers and colleagues to co-construct their knowledge and awareness and share their thoughts to discover solutions in a supportive and safe environment (Ahn & Kwon, 2025; Aşık & Gönen, 2016). In this way, CRP, "a more dialogic, data-led and collaborative approach to reflective

practice" (Mann & Walsh, 2013, p. 291), is implemented to involve pre-service EFL teachers to discuss and share ideas, experience, and theories in a collaborative learning environment to expand their knowledge and transform their practice.

CRP integrates collaboration and interaction with individual reflective practices in a team- or peer-based professional development. Supported by socio-constructivism (collaborative learning) and self-regulated learning (reflective learning) (Butler & Cartier, 2004), preservice EFL teachers can work with their colleagues to actively reflect socially on their practice, examine and discover teaching collectively, gain scaffolding and feedback in a team, complete collaborative projects, and observe, question, and change principles and practice collaboratively (Mitchell & Sackney, 2011). Thus, pre-service teachers experience a peer coaching and team teaching approach with which they can build collegiality by working with others socially and reflectively (Richards & Farrell, 2005). Several scholars have proposed that the effectiveness of collaborative reflection in fostering reflection among preservice teachers relies on two key factors: the guidance provided by knowledgeable individuals, such as teacher educators and peers, and the social organization of the collaboration, namely, team teaching or observation (Ghamoushi, 2025; Moore-Russo & Wilsey, 2014). However, as CRP is a time-consuming, dynamic, and challenging process and needs interaction among individuals, preservice teachers may find it hard to approach teacher educators and peers to collaborate with. Also, as Vangrieken et al. (2015) noted, teachers may encounter a wide range of challenges and issues when it comes to collaboration. These include social competitiveness, personal conflicts, limited autonomy, growing tensions, excessive workloads, and the need to conform. Hence, AI-based ChatGPT can act as a knowledgeable source for their collective inquiry during CRPs with an AI tool that can be collaborative, approachable, reliable, supportive, and enriched. ChatGPT can support pre-service teachers spontaneously, provide feedback, examine performance, improve knowledge and skill, and modify teaching beliefs and practice (Pan et al., 2025). Therefore, this study aims at exploring how preservice EFL teachers experience using AI-based ChatGPT to conduct CRP to improve more practically and actively through an AI system during language teacher education program.

2.2. Artificial intelligence in teacher education

In today's rapidly changing world, where information, knowledge, and skills are constantly being updated, educators require effective tools to assist them in their teaching endeavors to enhance the quality of education. Given this context, AI has emerged as a promising remedy for teachers to update their knowledge and practice, tackle their obstacles, and amplify their productivity (Luan et al., 2020). One recently-used AI tool that has received consideration is ChatGPT, created by OpenAI, to enable the production of written content by utilizing user prompts and generating intelligent and relevant answers to user queries (Moorhouse, 2023). Hence, ChatGPT, as a new technological innovation, is capable of reforming and transforming teacher education by engaging preservice teachers in an AI-based environment to receive immediate feedback on their first teaching practices, make personalized learning and assistance available for their self-directed learning practices, meet the needs and interests of each pre-service teacher's individualized learning style, and clarify key and abstract concepts for them during reading (Kuleto et al., 2021).

ChatGPT, therefore, has the ability to function as a virtual tutor, providing additional educational assistance outside of the traditional classroom-oriented teacher education. During preservice teachers' practicum and when they become novice teachers, ChatGPT has the capability to support curriculum and material development through the provision of pertinent content, propose effective teaching approaches and authentic tasks, and facilitate lesson planning (Baidoo-Anu & Owusu Ansah, 2023). Furthermore, it empowers teachers to identify motivating and appropriate content, therby supporting their instructional and subject-matter knowledge. Dwivedi et al. (2023), for instance, confirmed that ChatGPT can facilitate the process of grading, recognize errors and gaps, and design plans. In addition, it can provide pre-service teachers with continuous professional development practices, promoting a culture of ongoing enhancement.

The use of ChatGPT serves many other benefits in teacher education as well. ChatGPT, for instance, can help with managing administrative responsibilities that can be messy for preservice teachers, such as arranging schedules, handling administrative queries, and facilitating communication with parents (Bridgeman & Shipman, 2022). This empowers preservice teachers to dedicate more time to instructional activities. Accordingly, Kuhail et al. (2022) found that ChatGPT can serve as teaching agents, peer agents, and motivational agents. This intelligent system can offer personalized and interactive experiences for preservice teachers during teacher education (Li et al., 2019), fostering a human-centered and personalized instructional approach. In this case, ChatGPT can function as a conversational agent by simulating a human-like and real-world interaction (Zeng et al., 2024). Given a customized learning experience, ChatGPT facilitates personalized learning, catering to individual learning styles, pace, motivation, and engagement. They encompass a wide range of meaningful and purposeful activities, effective classroom management techniques, and procedural guidelines to teach initially, which can enhance the quality of teacher education (Ali et al., 2023; Rahman & Watanobe, 2023).

Although ChatGPT has offered several advantages, it may negatively affect preservice teachers' practices. According to Halleem et al. (2022), the establishment of strong relationships between preservice teachers and their peers and teacher educators is crucial for professional teacher education. However, an excessive dependence on AI without a critical and creative view could diminish the human-oriented interactions and mentorship that teacher educators provide, potentially affecting pre-service teachers' creative, emotional, and social development (Morron, 2023). Furthermore, preservice teachers and teacher educators may require additional training to effectively integrate AI into their teaching methods, since the rapid advancement of technology can create skill gaps and make it challenging for them (Luckin & Holmes, 2016). It is perceived that preservice EFL teachers can implement CRP with ChatGPT to have social interactions, reflect on and examine their teaching practices, acquire and expand their practical knowledge, receive help and scaffolding, possess a continued cooperative work in the ongoing professional development process, and help whenever they need. The current study probes into the following research questions to achieve the goals of the study:

1. How can pre-service EFL teachers implement CRP by the means of ChatGPT to develop professionally?
2. What are the benefits and challenges of doing CRP with ChatGPT for preservice EFL teachers?

3. Methodology

3.1. Research design

In the current qualitative study, preservice EFL teachers and teacher educators were explored to see how they perceive the use of AI to do CRP during their teacher education programs and explore the benefits and challenges of conducting CRP with ChatGPT. The present study followed an interpretive qualitative design to rich and contextualize data from the participants (Ary et al., 2019). An interpretive investigation is used to obtain detailed explanations aimed at comprehending and describing a phenomenon (AI-based CRP during teacher education) by utilizing data through various methods, including interviews, narratives, observations, and group discussions. In this basic interpretive study, the phenomenon (AI-based CRP with ChatGPT) and its processes were described to show novel ideas, practices, perceptions, and experiences (Creswell & Poth, 2016).

3.2. Participants and setting

To conduct an interpretive qualitative study, eight pre-service EFL teachers and two teacher educators were selected purposefully from a public teacher education university in Iran. The university encompasses a vigorous English department, providing bachelor of arts, master, and PhD degrees in language teaching. This internationally renowned university, recognized for its interdisciplinary approach, places excessive prominence on integrating technology into educational practices. It has a strong language and computer laboratory, a free access to fast-speed internet connection, online free courses and programs for the students and faculties, established synchronous and asynchronous online learning platforms, and a keen interest on using technology and AI for educational procedures. Moreover, it offers a variety of courses, including computer-assisted language learning, digital literacy, online assessment and e-learning, and other technology-oriented learning programs. Therefore, the university was open to AI-based practices to enhance teacher education programs.

Pre-service EFL teachers were on the first (60%) and second year (40%) of their Bachelor of arts in teaching English as a foreign language (TEFL) program. The participants were selected purposively, and the sample includes subjects who are regarded as representative of the population, utilizing a nonprobability sampling technique (Obilor, 2023). The researchers had a short interview with pre-service teachers for 15 minutes and asked some relevant questions to select those who were engaged in teaching at public and private schools and language institutes, had a decent level of computer literacy and technological devices, and became motivated to take part in research that requires time and effort were. Eight Iranian pre-service EFL teachers (females = 3; males = 5; Mean age = 22; SD = 3.75) were selected for the study. They were aged between 19 and 26, with 1 to 2 years of teaching experience.

The selection of eight preservice EFL teachers was motivated, following the data saturation, a crucial criterion in qualitative research. Data saturation occurs when no new themes or understandings appear from the data, and new participants add no new information to represent enough depth and breadth to solve the research questions (Braun & Clarke, 2021). In the present study, data were started through collecting from eight teachers, and the ninth teacher provided no new information. Moreover, two teacher educators were selected as they had a strong research and teaching background in teaching with computer and AI within a teacher education context. Both were male and hold PhD in applied linguistics. Also, they were an assistant professor and a lecturer, with five to 10 years of teaching experience, and they have taught in different language institutes, schools, and universities.

3.3. Instruments

In this study, the researchers used four major instruments to collect data as shown and briefly explained in Table 1. We used four different instruments to triangulate by instrumentation to help collect various data types for thick description and validity purposes, as suggested by Serafini et al. (2015). More detailed information regarding these instruments is presented in the following section.

Table 1. Instruments used for the collection of data in the present study

Narratives	Preservice teachers wrote three narratives (each around 1,000 words) about their experiences with CRP and ChatGPT.
Nonparticipant observations	Teacher educators observed preservice teachers' classes and wrote recorded their observations about preservice teachers' development.
Group discussions	Audio recordings and field notes were taken during weekly group discussions.
Unstructured interviews	Open-ended interviews were conducted with preservice teachers and teacher educators, and were audio-recorded for transcription.

3.4. Data collection procedure

After selecting pre-service EFL teachers and teacher educators purposively, the first researcher conducted three one-hour workshops about CRP and AI-based ChatGPT for teacher educators and pre-service teachers. During the first session, the researcher discussed the definitions, models, and procedures of doing CRP, such as how to plan, act, observe, and reflect collectively. For the second session, the participants were informed about the significance of AI and ChatGPT to improve the quality of their teacher education programs by reflecting and collaborating with ChatGPT. Last but not least, a question-and-answer part took place for an hour to address any queries related to procedures and models of CRP and ChatGPT (e.g., Do you have any concerns or issues regarding the implementation of CRP?, Do you need more information about the procedures?). Consequently, the participants were instructed to utilize and apply ChatGPT while encouraging pre-service EFL teachers to engage in CRPs. The participants were expected to conduct CRP via ChatGPT with the help of their teacher educators for three months.

During the study, preservice teachers had to conduct CRP with ChatGPT for reflecting on their teaching practices, identify problems and find solutions, design tasks and activities, prepare the lesson plans, find useful resources, receive feedback, and reform knowledge and practice. The first researcher was approachable to provide continuous support, recommendation, and feedback during their CRP. To gather data, they were encouraged to write three ***narratives*** in English (each should be around 1,000 words) to indicate how they could develop their teaching skills, knowledge, and attitudes through conducting CRP with ChatGPT. They had to write about the process, procedure, and experience of implementing CRP with ChatGPT for enhancing their teaching practice. All pre-service teachers completed the three narratives, except one who was busy and provided only one narrative. The researcher gathered their narratives at the end of the study. Besides, teacher educators had to ***observe*** student teachers' classes for three separate sessions: once at the beginning, once in the middle, and once at the conclusion of their teaching period. This was done to assess how student teachers could develop their knowledge, skill, and practice by doing CRP with ChatGPT. Their observation was unstructured without having a framework to understand their improvement holistically. Teacher educators had to write a narrative to show how student teachers have developed, what positive changes felt, and what the benefits and challenges were. The narratives were gathered at the end of the study.

They had ***group discussions*** to receive guidance, support, knowledge, and feedback from their teacher educators personally and virtually through social media applications and interact with other peers. At the end of each week, they had a meeting for an hour and half to talk about their AI-based CRP practices, perceptions, experiences, and processes with peers and teacher educator in a group to know others' experience, learn from them, possess scaffolding, have a mutual interaction, and develop collectively. The teacher educators and the lead researcher moderated and facilitated the process of group discussions. Hence, the lead researcher observed their group discussion practices without active participation. The researcher collected the data through field notes and audio recordings. In addition, the teacher educators and pre-service teachers were interviewed for 45 minutes through an ***unstructured interview*** with some open-ended questions related to the role of ChatGPT in CRP.

To make sure of the content validity of the interview questions, they were reviewed by one expert in applied linguistics familiar with conducting interviews, pilot-tested with some participants, and refined based on feedback received from them. Moreover, data triangulation with narratives, observations, and group discussions further supported the credibility of the findings. Therefore, the questions could be modified during the interview to know how CRP could be developed though ChatGPT, the benefits and disadvantages, and the opportunities and challenges of this mode of AI-based teacher education. During the interviews, teacher educators share their perceptions of how pre-service teachers could develop professionally through this model. All interviews were audio recorded for further transcription and analysis.

3.5. Data analysis

The collected data included narratives, audio recordings, and field notes gained from interviews, narratives, observations, and group discussions. These materials were transcribed by the researchers to enhance analysis and thematic categorization to answer both research questions. Inductive thematic analysis (Braun & Clarke, 2006) was used to reach themes organically from the rich, varied data, producing a flexible, data-driven, and rigorous investigation of participants' perceptions and experiences. The researchers accurately reviewed the data multiple times to recognize key concepts, meanings, and ideas, which were then coded. Extracts were systematically categorized, with sub-categories grouped under overarching themes.

To guarantee the robustness of the coding process, inter-rater reliability was calculated. As Krippendorff (2004) highlights, inter-coder reliability necessitates that two raters assign the same code to the same unit of text. In this study, the first researcher coded all the data, and a Ph.D. holder in applied linguistics was invited to independently analyze a 25% of data, as is common in qualitative studies (e.g., O'Connor & Joffe 2020) and the inter-rater reliability score was computed to assess the consistency of coding. The results of inter-coder reliability, using Cronbach's alpha, was .91, showing a relatively high consistency between the two coders. Following Modarresi (2025), in cases of disagreement between the two coders, a third coder was invited. The process of coding was repeated to resolve discrepancies through negation and discussion until full agreement was reached. We also used rechecking transcripts, refining codes, and validating categorizations, as suggested by Ary et al. (2019), to increase the reliability and validity of the findings.

4. Results

In this part, we present the results of the study. For clarity of organization of the findings, we state them research question by question. For each research question, the major themes are presented, followed by subthemes and relevant extracts from the data collected from the instruments discussed in the previous section. The source of quotations is identified as follows: observation (O), group discussion (GD), interviews (I), and narratives (N).

4.1. Investigation of the first research question: How can preservice EFL teachers implement CRP by the means of ChatGPT to develop professionally?

Table 2 outlines the major themes and subthemes for the first research questions, and they are elaborated on, followed by extracts from the participants, in the following paragraphs.

Table 2. Themes and subthemes related to the first research question

Major Themes	Subthemes
AI's potential to instill creativity	Enhances experiential learning, autonomy, collaboration, and innovation. Provides real-world content, feedback, and problem-solving. Supports lesson planning, task design, and teaching techniques.
AI's capability for professional and autonomous development	Encourages reflection, collaboration, and practical experience Connects theory and practice. Provides constructive feedback and scaffolding. Enhances self-regulation and emotional control.

4.1.1. AI has the potential to instill creativity into preservice teachers

AI was an innovative, interactive, and effective partner for preservice teachers, facilitating their teaching, learning, and professional development. This innovation was mainly due to its "*accessibility to current and first-hand knowledge of teaching theories and practices"* as stated by preservice teacher 8 (O). In previous traditional teacher education, teacher educators explained the abstract concepts and pre-service teachers absorbed information, while "*AI could enhance pre-service teachers' experiential learning, autonomy, collaboration skills, transformation, and innovation"* (Teacher Educator 2[I]). Another teacher educator observed that: "*Three student teachers could improve their experiential learning while participating actively, receiving real-world and relevant content and feedback, reflecting on their practices with ChatGPT, having collaboration, solving problems, and creating safe learning environment to improve their teaching practices continuously"* (O). AI tools were used in a creative way to support teacher educators and preservice teachers in order to gain authentic and real-world content, work as a teaching assistant to teach and assess, enhance their interactive practices, and improve their professional development, provide immediate, constructive, and supportive feedback for pre-service teachers during their initial teaching practices, as stated preservice Teacher 2 as follows:

> Umm … When I used ChatGPT to enhance my teaching, its feedback encouraged me to consider different factors, such as students' learning styles, contextual factors, facilities, students' motivation and proficiency, policies and initiatives, and so more. It was such a big help related to my personal and professional needs. (I)

ChatGPT was also seen as a personalized learning system that could help with designing tasks and lesson plans, observing teaching quality, checking the progress, grading assessment, and managing responsibilities. Preservice teacher 8, for instance, pointed out that "*I used ChatGPT to meet my needs and interests. Besides, it helped me with my teaching from the beginning to the end, such as devising lesson plans, preparing tasks, and suggesting teaching techniques"* (GD). ChatGPT could provide the required knowledge and skills, assistance and help, and solutions and ideas to problems for pre-service teachers. As stated by teacher educator 1, "*I am sure that pre-service teachers could be more knowledgeable, professional, and effective by the means of AI, since AI was such a helpful partner to them. They could learn how to teach skills such as reading or listening and speaking or writing more effectively by interacting with ChatGPT"* (N). Furthermore, preservice teachers could gain confidence to control, manage, and direct their own learning autonomously by receiving constant guidance and support from an AI-based ChatGPT, as explained by preservice teacher 7: "*I learned how to teach grammar more effectively by planning lessons, preparing authentic content, designing tasks, administering through games, and assessing creatively with the help of ChatGPT".* (I).

4.1.2. Doing CRP Through ChatGPT leads to professional and autonomous development

Pre-service teachers were able to collaborate and reflect on their experiences, search for new information, expand their knowledge and awareness, and generate innovative ideas and actions resourcefully with ChatGPT, as stated by preservice teacher 5: "*I could reflect on my vocabulary teaching, and knew that students' previous learning, their motivation and interest, and learning styles*

should be considered, so collaborating with ChatGPT enhanced my awareness regarding teaching vocabulary" (GD). Furthermore, doing CRP with ChatGPT helped preservice teachers monitor their teaching, gain practical experience, find the gaps, conduct strategies wisely, and connect theories and practice professionally. Teacher educator 1, for example, asserted that "*Pre-service teachers could do the same procedure of human-based CRP with AI-based ChatGPT, such as observing, planning, acting, and reflecting"* (N). Personally, they exercised more agency, took wise actions, changed their belief systems, controlled and managed their emotions, and became more creative and critical. In this regard, preservice teacher 1 expressed the following ideas:

> *"Well, as a novice teacher, I cannot regulate my emotions perfectly, specifically the negative ones. But, ChatGPT enhanced my cognition and awareness to control and manage it. Considering the time when I was anxious of students' poor behavior, ChatGPT provided some solutions to implement them in my class, and I was successful."* (I)

When CRP was conducted with ChatGPT, preservice teachers received more constructive feedback, scaffolding, and support for their teaching practice and pedagogical choices during and after teaching hours. To give an example, teacher educator 2 explained that preservice teachers could "*examine practice and recognize and solve problems, adapt and change the established practices that were not effective for present students, and foster situated awareness of the context and students"* (I). Accordingly, pre-service teachers became more analytical and reflective, possessed a collective inquiry, changed for positive results, adapted their thoughts and actions, and shared thoughts and emotions with ChatGPT. In this way, as teacher educator 1 observed, "*pre-service teachers turned out to be more self-regulated and collaborative to develop professionally with AI continuously"* (O)"

4.2. Investigation of the second research question: What are the benefits and challenges of doing CRP with ChatGPT for preservice EFL teachers?

The major themes and subthemes regarding the second research question are presented in Table 3. Detailed explanations accompanied by the participants' quotes are discussed in what follows.

Table 3. Themes and subthemes related to the second research question

Major themes	Subthemes
Advantages of Using AI for preservice teachers	Accessibility for use anywhere and anytime Personalized learning and teaching assistance Provision of constructive feedback and self-directed learning Enhancement of engagement, motivation, and self-regulation
Challenges of Using AI for doing CPR for preservice teachers	Limited interpersonal and human-oriented relationships Restricted social, emotional, and cultural development Limited interaction with peers and colleagues Frustration is use of AI due to poor internet connectivity Restricted agency and innovation as result of overreliance on AI

4.2.1. Advantages of using AI in teacher education

AI-based ChatGPT can be an enriched source for pre-service teachers and teacher educators to find related articles and books, create tasks and exercises, reflect on experiences, examine performance, and recognize gaps of knowledge. Preservice 3, for example, stated that "*ChatGPT guided my reflection on my classroom behavior, attitudes, and discourse constructively and offered some recommendation to enhance it"* (N). ChatGPT is also approachable anytime and anywhere with which teachers can use comfortably to upgrade their knowledge, skill, and practice. Moreover, it is seen as a personalized learning and teaching assistance by having constructive feedback, learn from their own experience, motivate a self-directed learning approach, and fulfill their needs and interests, as pointed out by preservice teacher 1: "*Such an awesome tool! I used it to make some concepts clear, such as what self-correction is and how it should be used in my teaching"* (I).

Additionally, pre-service teachers can receive help regarding how to prepare and generate authentic, relevant, and beneficial content for students, plan the way they should present them with appropriate and engaging teaching approaches and techniques, and use different forms of assessment to provide feedback and boost their practices formatively. In this regard, preservice teacher 2 offered the following ideas: *"So, we enhanced our engagement and motivation, self-regulation and self-development, and knowledge and skill by reflecting collaboratively with ChatGPT. The result led to more effective materials, content, resources, teaching, and evaluation"* (N).

4.2.2. Challenges preservice teachers face in using ChatGPT to do CRP

There are several negative points for doing CRP with ChatGPT, such as lack of inter-personal and human-oriented relationships in a group, team, or a community to develop mutually with peers and colleagues. For example, teacher educator 2 noted that "*I guess that … pre-service teachers could work together with team teaching and coaching techniques to break their isolation" (*N). Besides, teachers focused on the input received from ChatGPT with little critical or creative thinking, which can damage

teachers' agency and innovation. "*I only listened to whatever ChatGPT offered, and I confirmed whatever presented",* preservice teacher 4 mentioned in one of the group discussions. Additionally, pre-service teachers should be developed socially, emotionally, and creatively through social practices, since they have restricted interaction with others while doing CRP with ChatGPT. This is what preservice teacher 5 lamented about: "*I could interact with it, yet there was no fun, interaction, feeling, and encouragement"* (I).

As internet connection can sometimes be poor in Iran, there are some connection problems during the connection, which can frustrate users. Preservice teachers and teacher educators should be trained to gain AI literacy and use ChatGPT effectively, influencing the way they implement AI for learning and teaching purposes. Also, CRP with ChatGPT limits the cultural awareness and increases unethical behaviors and practices, such as overreliance on ChatGPT. Preservice teacher 7 highlighted this challenge as follows: "*ChatGPT made me lazy as I used it for all of my teaching practices. Also, the poor internet connection was an obstacle and hinder*" (N). To enhance their contextual and cultural understanding, teachers should interact with other peers and teachers. Moreover, CRP can be conducted in a community of practice with other peers to create a friendly and supportive learning atmosphere. However, this interaction is unsatisfactory, as preservice teacher 4 explains: "*In my opinion, since I could not connect well due to limited speed of internet, have social connection, and possess little knowledge of AI, I faced many challenges throughout the process"* (I).

5. Discussion

As far as the first research questions is concerned, the findings of this study show that pre-service EFL teachers can use CRP with ChatGPT to grow professionally. ChatGPT functions as a teaching assistant and a virtual tutor, providing advanced and innovative chances for pre-service teachers to improve their instructional beliefs, theories, principles, and techniques. By means of ChatGPT, pre-service teachers can get the latest teaching knowledge, methods, and techniques, nurturing their instructional practices and experiential learning (Jeon & Lee, 2023; Kohnke et al., 2023). The results of this study support previous research, which underscores the role of ChatGPT in boosting professional development. To illustrate, Kasneci et al. (2023) found that AI tools like ChatGPT offer direct and practical feedback, empowering pre-service teachers to improve their theoretical knowledge and practical skills. Likewise, Koedinger et al. (2012) highlighted that AI-based systems can link theory and practice by proposing systematic support and assistance. In the present study, pre-service teachers stated that ChatGPT helped them reflect collectively, resolve instructional issues, and foster their pedagogical competence and teaching skills.

Although the findings of the present study confirm the professional benefits of ChatGPT, the findings contrasts with some of previous studies. As such, Bridgeman and Shipman (2022) discussed that AI tools may require the depth of human interaction desired for complex professional development. This study recognizes that although ChatGPT is effective in offering personalized feedback and tasks, it cannot completely replicate the social and emotional aspects of peer collaboration, which are vital for complete professional development.

Concerning the second research question, the study recognized numerous benefits of employing CRP with ChatGPT for pre-service EFL teachers. Initially, ChatGPT offers genuine and enhanced input, enhancing its access to present language teaching knowledge, developing educational practices and experiential learning (Dwivedi et al., 2023). Then, it affords instantaneous, constructive, and procedural feedback, which certainly impacts pre-service teachers' social interactions, reflective learning, and teaching practices (Kasneci et al., 2023). Furthermore, ChatGPT acts as a personalized learning system, producing tasks and lesson plans suitable for specific needs and interests, thus boosting instructional quality and efficiency. These findings confirm some of the findings from previous studies conducted recently. For instance, Lim et al. (2023) highlighted the ChatGPT's capacity to nurture autonomy and self-regulation among learners, which supports the observation of the study that pre-service teachers developed more autonomy, creativity, and motivation. Furthermore, Morron (2023) highlighted the role of ChatGPT in supporting teachers socially and emotionally, which ChatGPT's potential to monitor and increase emotional and behavioral features of teaching.

In spite of these benefits, the study also acknowledged challenges connected with employing ChatGPT for CRP. One major restriction is the reduced interpersonal and peer-based communications, which adversely influence emotional, social, and cultural development. Even though ChatGPT is collaborative, it cannot replicate the empathetic feedback, game-based activities, and fun-making moments that peer collaboration suggests (Jeon & Lee, 2023). This finding is in contrast with that of Kohnke et al. (2023), who contended that AI tools can completely support social learning. However, this study stresses that peer collaboration becomes indispensable for overcoming separation and nurturing original teaching practices. Another challenge is the potential over-reliance on ChatGPT, which may impede the growth of critical thinking and creative reflections. This supports the assertion expressed by Kasneci et al. (2023), who concluded that too much reliance on AI tools could damage teachers' autonomous decisions and creative actions. Besides, the study underlines real-world challenges, including the necessity for fast internet connection, better AI literacy among teachers, and ChatGPT's restricted understanding of various cultural and contextual features. These restrictions imply that CRP with ChatGPT should be complemented with peer and community-based practices to produce a supportive learning setting.

To recapitulate, the findings of this study both support and deviate from previous research. On the one hand, the study supports the role of ChatGPT in augmenting professional development, as maintained by Kasneci et al. (2023) and Koedinger et al. (2012). On the other hand, it underscores restrictions, including reduced peer collaboration and too much dependence on AI, which differ with the more positive views of Kohnke et al. (2023). Also, the study highlights the necessity for a stable approach that incorporates AI tools with human collaboration, a perspective that is progressively highlighted in current literature (e.g., Lim

et al., 2023; Morron, 2023). Hence, this study indicates that pre-service EFL teachers can conduct CRP with ChatGPT to develop professionally, benefiting from personalized feedback, enhanced input, and experiential learning. However, the challenges such as incomplete peer interaction and over-reliance on AI must be addressed.

6. Conclusion

This study investigated the potential of incorporating ChatGPT as an AI-based tool to foster CRP among pre-service EFL teachers. Narratives, interviews, observations, and group discussions were used to explore the opinions of eight pre-service EFL teachers and two teacher educators. The findings revealed that ChatGPT can function as an accessible, creative, and educational instrument to support pre-service teachers in their professional development. Specifically, it can act as a virtual mentor and collaborator, empowering pre-service teachers to reflect on novel experiences, obtain knowledge, solve challenges, plan strategic actions, and nurture creativity. As such, ChatGPT's capacity to offer instant, constructive, and helpful feedback during initial teaching practices underlines its value as a resource for novice teachers. Although we confirm the relative utility of AI tools such as ChatGPT in teacher education programs to foster preservice teachers' professional development, teacher educators need to consider the limitations AI tools have, and such tools should not be considered as panacea for all the problems teachers may face in their future careers.

The present study has its own limitations. Firstly, the small sample size hampers the generalizability of the findings, signifying the necessity for future research with a larger and more varied participant. Additionally, the study's qualitative design, could be added by a mixed-methods approach to offer a more comprehensive understanding of how AI tools such as ChatGPT influence pre-service teachers' CRP development. As an example, a pre-test-post-test experimental design can quantitatively measure variations in pre-service teachers' attitudes, behaviors, and practices, while qualitative methods can explore their reflective and collaborative experiences more deeply. Furthermore, future studies should consider examining how pre-service EFL teachers from various cultural contexts and teacher education programs may apply AI tools for reflective and collaborative development, with a focus on their attitudes, emotions, and teaching strategies. Such inquiries will offer a more detailed understanding of the role of AI in teacher education.

Despite the limitations, the findings may offer implications for teacher educators, school administrators, stakeholders, and policymakers to integrate AI tools like ChatGPT into teacher education programs. In this case, they can improve pre-service teachers' cognitive and social functioning, nurturing a generation of educators who are not only technologically skillful but also culturally approachable and reflective in their teaching practices. Eventually, the thoughtful incorporation of AI in teacher education has the potential to renovate how pre-service teachers develop professionally and autonomously, making them ready for the challenges the language classes may pose in the future.

7. References

Ahn, J., & Kwon, J. (2025). How do we build and sustain a teacher community? The case of Korean language teachers. *Asia Pacific Journal of Education*, 1-15. https://doi.org/10.1080/02188791.2024.2449133

Ali, J. K. M., Shamsan, M. A. A., Hezam, T. A., & Mohammed, A. A. (2023). Impact of ChatGPT on learning motivation: Teachers' and students' voices. *Journal of English Studies in Arabia Felix*, *2*(1), 41–49. https://doi.org/10.56540/jesaf.v2i1.51

Alvarado Gutiérrez, M., A Neira Adasme, M., & Westmacott, A. (2019). Collaborative reflective practice: Its influence on pre-service EFL teachers' emerging professional identities. *Iranian Journal of Language Teaching Research, 7*(3), 53–70. https://doi.org/10.30466/ijltr.2019.120736

Ary, D., Jacobs, L. C., Irvine, C. K. S., & Walker, D. (2019). *Introduction to research in education* (10th ed.). Cengage Learning.

Aşık, A., & Kuru Gönen, S. İ. (2016). Pre-service EFL teachers' reported perceptions of their development through SETT experience. *Classroom Discourse*, *7*(2), 164–183. https://doi.org/10.1080/19463014.2016.1150865

Baidoo-Anu, D., & Ansah, L. O. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, *7*(1), 52–62. https://doi.org/10.61969/jai.1337500

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher*, *33*(8),3–15. https://doi.org/10.3102/0013189X033008003

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, *3*(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2021). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative research in Sport, Exercise and Health*, *13*(2), 201–216.

https://doi.org/10.1080/2159676X.2019.1704846

Bridgeman, A. J., & Shipman, M. A. (2022). Artificial intelligence and education: Promises and perils. *Journal of Educational Technology Systems, 49*(1), 3–22. http://dx.doi.org/10.1177/0047239520984205

Butler, D. L., & Cartier, S. C. (2004). Promoting effective task interpretation as an important work habit: A key to successful teaching and learning. *Teachers College Record*, *106*(9), 1729–1758. https://doi.org/10.1111/j.1467-9620.2004.00403.x

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, Article e107468. https://doi.org/10.1016/j.chb.2022.107468

Celik, I., Dindar, M., Muukkonen, H., & Järvelä, S. (2022). The promises and challenges of artificial intelligence for teachers: A systematic review of research. *TechTrends*, *66*(4), 616–630. https://doi.org/10.1007/s11528-022-00715-y

Cirocki, A., & Farrell, T. S. (2017). Reflective practice for professional development of TESOL practitioners. *The European Journal of Applied Linguistics and TEFL*, *6*(2), 5–23.

Cole, D. C., Parkes, M. W., Saint-Charles, J., Gislason, M., McKellar, K., & Webb, J. (2018). Evolution of capacity strengthening: insights from the Canadian community of practice in ecosystem approaches to health. *Transformative Dialogues: Teaching and Learning Journal*, *11*(2), 1-21.

Creswell, J. W., & Poth, C. N. (2016). *Qualitative inquiry and research design: Choosing among five approaches*. Sage publications.

Daniel, G. R., Auhl, G., & Hastings, W. (2013). Collaborative feedback and reflection for professional growth: Preparing first-year pre-service teachers for participation in the community of practice. *Asia-Pacific Journal of Teacher Education*, *41*(2), 159–172. https://doi.org/10.1080/1359866X.2013.777025

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., & Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, *71*, Article e102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Farrell, T. S. (2012). Novice-service language teacher development: Bridging the gap between preservice and in-service education and development. *TESOL Quarterly*, *46*(3), 435–449. https://doi.org/10.1002/tesq.36

Farrell, T. S. (2015). *Reflective language teaching: From research to practice*. Bloomsbury Publishing.

Farrell, T. S. (2022). I felt a sense of panic, disorientation and frustration all at the same time: the important role of emotions in reflective practice. *Reflective Practice*, *23*(3), 382–393. https://doi.org/10.1080/14623943.2022.2038125

Ghamoushi, M. (2025). The impact of collaborative reflection and dialogic journals on EFL teachers' ecological agency. *Reflective Practice*, *26*(1), 1–24. https://doi.org/10.1080/14623943.2024.2412872

Gudeta, D. (2022). Professional development through reflective practice: The case of Addis Ababa secondary school EFL in-service teachers. *Cogent Education*, *9*(1), Article e2030076. https://doi.org/10.1080/2331186X.2022.2030076

Guskey, T. R. (2002). Professional development and teacher change. *Teachers and Teaching*, *8*(3), 381–391. https://doi.org/10.1080/135406002100000512

Gutiérrez, M. V. A., Adasme, M. A. N., & Westmacott, A. (2019). Collaborative reflective practice: Its influence on preservice EFL teachers' emerging professional identities. *Iranian Journal of Language Teaching Research*, *7*(3), 53–70. https://doi.org/10.30466/ijltr.2019.120736

Haleem, A., Javaid, M., Qadri, M. A., & Suman, R. (2022). Understanding the role of digital technologies in education: A review. *Sustainable Operations and Computers*, *3*, 275–285. https://doi.org/10.1016/j.susoc.2022.05.004

Javaid, M., Haleem, A., Singh, R. P., Khan, S., & Khan, I. H. (2023). Unlocking the opportunities through ChatGPT tool towards ameliorating the education system. *BenchCouncil Transactions on Benchmarks, Standards and Evaluations*, *3*(2), Article e100115. https://doi.org/10.1016/j.tbench.2023.100115

Jeon, J., & Lee, S. (2023). Large language models in education: A focus on the complementary relationship between human teachers and ChatGPT. *Education and Information Technologies, 28*(12)1–20. https://doi.org/10.1007/s10639-023-11834-1

Jia, C., Hew, K. F., Bai, S., & Huang, W. (2022). Adaptation of a conventional flipped course to an online flipped format during the Covid-19 pandemic: Student learning performance and engagement. *Journal of research on technology in education*, *54*(2), 281–301. https://doi.org/10.1080/15391523.2020.1847220

Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, *103*, Article e102274. https://doi.org/10.1016/j.lindif.2023.102274

Kim, D., & Lee, S. (2002). Designing collaborative reflection supporting tools in e-project-based learning environments. *Journal of Interactive Learning Research*, *13*(4), 375–392.

Kim, J. (2023). Leading teachers' perspective on teacher-AI collaboration in education. *Education and Information Technologies, 29*(7), 1–32. https://doi.org/10.1007/s10639-023-12109-5

Koedinger, K. R., Corbett, A. T., & Perfetti, C. (2012). The knowledge-learning-instruction framework: Bridging the science-practice chasm to enhance robust student learning. *Cognitive science*, *36*(5), 757–798. https://doi.org/10.1111/j.15516709.2012.01245.x

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 535-550. https://doi.org/10.1177/00336882231162868

Krippendorff, K. (2004). *Content analysis: An introduction to its methodology* (2nd ed.). Sage.

Kuhail, M. A., Thomas, J., Alramlawi, S., Shah, S. J. H., & Thornquist, E. (2022). Interacting with a chatbot-based advising system: understanding the effect of chatbot personality and user gender on behavior. *Informatics, 9*(4), 1-24. https://doi.org/10.3390/informatics9040081

Kuleto, V., Ilić, M., Dumangiu, M., Ranković, M., Martins, O. M., Păun, D., & Mihoreanu, L. (2021). Exploring opportunities and challenges of artificial intelligence and machine learning in higher education institutions. *Sustainability*, *13*(18), Article e10424. https://doi.org/10.3390/su131810424

Lameras, P., & Arnab, S. (2021). Power to the teachers: an exploratory review on artificial intelligence in education. *Information*, *13*(1), 14. https://doi.org/10.3390/info13010014

Li, Y., Yang, H. H., & MacLeod, J. (2019). Preferences toward the constructivist smart classroom learning environment: Examining pre-service teachers' connectedness. *Interactive Learning Environments*, *27*(3), 349–362. https://doi.org/10.1080/10494820.2018.1474232

Lim, W. M., Gunasekara, A., Pallant, J. L., Pallant, J. I., & Pechenkina, E. (2023). Generative AI and the future of education: Ragnarök or reformation? A paradoxical perspective from management educators. *The International Journal of Management Education, 21*(2), Article e100790. https://doi.org/10.1016/j.ijme.2023.100790

Lin, C. C., Huang, A. Y., & Lu, O. H. (2023). Artificial intelligence in intelligent tutoring systems toward sustainable education: a systematic review. *Smart Learning Environments*, *10*(1), 41. https://doi.org/10.1186/s40561-023-00260-y

Luan, H., Geczy, P., Lai, H., Gobert, J., Yang, S. J., Ogata, H., ... & Tsai, C. C. (2020). Challenges and future directions of big data and artificial intelligence in education. *Frontiers in psychology*, *11*, Article e580820. https://doi.org/10.3389/fpsyg.2020.580820

Luckin, R., & Holmes, W. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson

Mann, S., & Walsh, S. (2013). RP or RIP: A critical perspective on reflective practice. *Applied Linguistics Review*, *4*(2), 291–315. https://doi.org/10.1515/applirev-2013-0013

Mann, S., & Walsh, S. (2017). *Reflective practice in English language teaching: Research-based principles and practices*. Taylor & Francis.

Meanwell, E., & Kleiner, S. (2014). The emotional experience of first-time teaching: Reflections from graduate instructors, 1997–2006. *Teaching Sociology*, *42*(1), 17–27. https://doi.org/10.1177/0092055X13508377

Mitchell, C., & Sackney, L. (2011). Building and leading within learning ecologies. In T., Townsend, & J., MacBeath (Eds.), *International handbook of leadership for learning* (pp. 975–990). Springer. https://doi.org/10.1007/978-94-007-1350-5_53

Modarresi, G. (2025). The role of motivational self-regulatory strategies in EFL learners' involvement in writing skills: A mixed-methods study. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills), 44*(1), 77-96. https://doi.org/10.22099/tesl.2024.51688.3351

Modarresi, G., & Alavi, S. M. (2014). Designing and validating a test battery of computerized dynamic assessment of grammar. *TELL, 14*(2), 1–29. https://doi.org/10.22132/tel.2014.53816

Moore-Russo, D. A., & Wilsey, J. N. (2014). Delving into the meaning of productive reflection: A study of future teachers' reflections on representations of teaching. *Teaching and Teacher Education, 37*, 76–90. https://doi.org/10.1016/j.tate.2013.10.002

Moorhouse, B. L. (2023). Virtual teaching abroad during initial teacher education: Pre-service teachers' professional learning. *Journal of Education for Teaching, 50*(3), 1–13. https://doi.org/10.1080/02607476.2023.2247351

Morron, M. (2023). *The impact of AI on learning and creativity: A critical look at the future.* Retrieved from https://www.linkedin.com/pulse/impact-ai-learning-creativity-critical-lookfuture-miriam-morron

O'Connor, C., & Joffe, H. (2020). Intercoder reliability in qualitative research: debates and practical guidelines. *International Journal of Qualitative Methods*, *19*. https://doi.org/10.1177/1609406919899220

Obilor, E. I. (2023). Convenience and purposive sampling techniques: Are they the same. *International Journal of Innovative Social & Science Education Research*, *11*(1), 1–7.

Pan, F., Zhao, L., & Cao, E. Y. (2025). Revisiting factors influencing strategies for enhancing pre-service teachers' digital competencies. *Education and Information Technologies*, 1-37. https://doi.org/10.1007/s10639-025-13416-9

Putman, M. S., Polly, D., & Fitzgerald, M. S. (2025). Moving from the periphery to practitioner: Investigating the development of a community of practice within a year-long clinical experience. *Journal of Teacher Education*. https://doi.org/10.1177/00224871251321898

Rahman, M. M., & Watanobe, Y. (2023). ChatGPT for education and research: Opportunities, threats, and strategies. *Applied Sciences*, *13*(9), 5783. https://doi.org/10.3390/app13095783

Richards, J. C., & Farrell, T. S. (2005). *Professional development for language teachers: Strategies for teacher learning*. Cambridge University Press.

Richards, J. C., & Lockhart, C. (1994). *Reflective teaching in second language classrooms*. Cambridge university press.

Rodgers, C. (2002). Defining reflection: Another look at John Dewey and reflective thinking. *Teachers College Record*, *104*(4), 842–866. https://doi.org/10.1111/1467-9620.00181

Salas-Pilco, S. Z., Xiao, K., & Hu, X. (2022). Artificial intelligence and learning analytics in teacher education: A systematic review. *Education Sciences*, *12*(8), 569. https://doi.org/10.3390/educsci12080569

Savoca, M., Dixon, K., & Zalewski, U. (2021). Building an ecosystem of diversity talent development through experiential learning. *Experiential Learning & Teaching in Higher Education*, *4*(2), 49–54.

Serafini, E. J., Lake, J. B., & Long, M. H. (2015). Needs analysis for specialized learner populations: Essential methodological improvements. *English for specific purposes, 40*, 11-26. https://doi.org/10.1016/j.esp.2015.05.002

Sweeting, A. M., & Carey, M. D. (2025). Strengthening the nexus of research and practice in CELTA pronunciation instruction. *System*, Article e103623. https://doi.org/10.1016/j.system.2025.103623

Vangrieken, K., Dochy, F., Raes, E., & Kyndt, E. (2015). Teacher collaboration: A systematic review. *Educational Research Review*, *15*, 17–40. https://doi.org/10.1016/j.edurev.2015.04.002

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes* (Vol. 86). Harvard university press.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., Jiang, Y., ... & Zhang, L. (2020). Deep learning in environmental remote sensing:

Achievements and challenges. *Remote Sensing of Environment*, *241*, Article e111716. https://doi.org/10.1016/j.rse.2020.111716

Zeichner, K., & Liu, K. Y. (2010). A critical analysis of reflection as a goal for teacher education. In N. Lyons (Ed.), *Handbook of reflection and reflective inquiry: Mapping a way of knowing for professional reflective inquiry*, (pp.67–84). https://doi.org/10.1007/978-0-387-85744-24

Zeng, Y., Rajasekharan, A., Padalkar, P., Basu, K., Arias, J., & Gupta, G. (2024). *Automated interactive domain-specific conversational agents that understand human dialogs*. Springer. https://doi.org/10.1007/978-3-031-52038-9_13

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 44-57

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.9157.1123

The Effect of Task-Specific Anxiety vs. Task-Specific Enjoyment on Language Mindset in L2 Listening Tasks

Amin Bigverdi[1*]
Abdorreza Tahriri[2]
Ayatollah Razmjoo[3]

[*1]Ph.D. Candidate in Applied Linguistics, Department of English Language and Literature, University of Guilan, Rasht, Iran
[2]Associate Professor of Applied Linguistics, Department of English Language and Literature, University of Guilan, Rasht, Iran
[3]Professor of Applied Linguistics, Department of Foreign Languages and Linguistics, Shiraz University, Shiraz, Iran

ARTICLE HISTORY

Received: 17 November 2024
Revised: 05 December 2024
Accepted: 27 January 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: a_bigverdi@webmail.guilan.ac.ir

ABSTRACT

This study addresses a significant knowledge gap in understanding the effects of task-specific enjoyment and anxiety conditions on the language mindset of Iranian EFL learners in a listening course. Using a quasi-experimental design, 75 male high school students were divided into three groups: a task enjoyment (TE) group, a task anxiety (TA) group, and a control group. The TE group participated in a low-stakes, autonomy-based listening task designed to foster enjoyment, while the TA group completed a high-stakes, evaluative listening task intended to induce anxiety. Data were collected using standardized instruments, including the Foreign Language Enjoyment Scale (FLES), the Foreign Language Classroom Anxiety Scale (FLCAS), and the Language Mindset Inventory (LMI). A series of T-tests and ANOVA were conducted to analyze pre- and post-test scores across the three groups. Results showed that the TE group experienced significantly higher levels of task enjoyment and subsequently developed significantly stronger growth-oriented beliefs regarding overall linguistic intelligence and second language learning compared to both the TA and control groups. Conversely, the TA group exhibited significantly higher anxiety levels and a shift toward more fixed language mindsets. No significant differences were found between groups regarding beliefs about age sensitivity in language learning. These findings highlight the importance of considering task-related emotions in language teaching and suggest that fostering enjoyment through autonomy-supportive tasks can promote a growth mindset, while anxiety-inducing tasks may reinforce fixed beliefs about language learning.

KEYWORDS: Language mindset; Listening comprehension; Task anxiety; Task enjoyment

1. Introduction

Task-Based Language Teaching (TBLT) is a teaching method that promotes the development of second language (L2) skills through meaning-focused tasks, giving individuals the opportunity to strengthen their communication abilities and engage in social interactions essential for their academic and professional goals (Ellis et al., 2019). In TBLT literature, tasks have been viewed either as static work plans, such as design and mode, created by material designers or as dynamic processes involving preparation, interaction, and repetition carried out by educators and students (Jackson, 2022; Samuda, 2015). Despite the structured nature of TBLT, a major challenge remains—learners vary in their ability to benefit from these tasks due to differences in cognitive resources and learning preferences. Even the most well-designed tasks do not guarantee equal learning outcomes for all students, as individual engagement and participation are essential in influencing success. Learners should be seen as active contributors to the educational process instead of just passive recipients, as their willingness to engage with tasks and their emotions significantly impact the effectiveness of TBLT (Almukhaild & King, 2023).

Cite this article: Bigverdi, A., Tahriri, A. and Razmjoo, A. (2025). The effect of task-specific anxiety vs. task-specific enjoyment on language mindset in L2 listening tasks. *Applied Linguistics Inquiry*, *3*(1), 44-57. doi: 10.22077/ali.2025.9157.1123

Emotions significantly influence how we learn, influencing how students engage with L2 tasks. Reeve (2024) describes emotions as short-lived reactions that support individuals in navigating challenges. While research has explored various emotions, much of the focus has been on anxiety (Dörnyei, 2009; Horwitz et al., 1986). Foreign Language classroom anxiety (FLCA), described as feelings of concern and negative responses associated with learning or using an L2 (MacIntyre, 1999), can negatively impact cognitive abilities, academic performance, self-efficacy, and willingness to communicate, sometimes even leading to language learning avoidance (Teimouri et al., 2019; Dewaele & MacIntyre, 2014). However, anxiety can also serve a positive role by fostering tension, focus, and resilience when kept at manageable levels (MacIntyre & Gregersen, 2012). According to Ellis (2003), anxiety becomes harmful when it leads to panic and impairs performance, but in moderate amounts, it can enhance motivation and engagement.

At the beginning of the new millennium, with the advent of positive psychology, there was a shift in focus within Western psychology towards highlighting human strengths and positive emotions, as opposed to solely concentrating on weaknesses (Seligman & Csikszentmihalyi, 2000). This shift has broadened the research focus to include both negative and positive emotions in language learning (Dewaele & MacIntyre, 2014; 2016; MacIntyre & Mercer, 2014). One commonly observed positive emotion is foreign language enjoyment (FLE), which plays a crucial role in shaping learners' experiences (Dewaele & MacIntyre, 2014). FLE is characterized by a combination of challenge and perceived ability, reflecting a learner's motivation to succeed despite difficulties (Dewaele & MacIntyre, 2016). A number of studies have indicated that there are noteworthy and favorable connections between FLE and perceived and actual foreign language ability, and academic achievements (Botes et al., 2020b; Dewaele et al., 2023; Teimouri et al., 2019). Of course, it can be challenging to examine how learners' emotions and their proficiency or learning are connected because the causal pathways between them can be bidirectional (Botes et al., 2020b).

Presenting emotions like enjoyment and anxiety as polar extremes on a continuum might give the impression of simplicity and clarity; however, studies have revealed that FLCA and FLE are actually related but independent, rather than being like a see-saw where they move up and down in opposite directions. In other words, having one emotion does not exclude the possibility of the other being there, and the reverse is also valid. FLE and FLCA are distinct emotions that do not oppose each other on a single continuum, and they are better seen as complementary to each other (Botes et al., 2022; MacIntyre & Ayers-Glassey, 2021). In recent years, research on FL learning and teaching has increasingly adopted a complex dynamic systems theory (CDST) perspective, emphasizing the ever-changing nature of emotions (MacIntyre et al., 2014). This approach suggests that emotions, though they may seem stable over long periods, actually fluctuate on much shorter timescales, such as seconds or minutes, and that learners within the same group do not necessarily experience emotions in the same way (Elahi Shirvan et al., 2020).

Given the significant role of emotions in shaping learning experiences, another key psychological factor influencing academic achievement is mindset. The idea of mindset is gaining traction within the realm of language education, and scholars are now exploring language mindsets, which are attitudes toward the malleability of language-learning abilities (Elahi Shirvan et al., 2021). This theory suggests that individuals have an inherent belief or mindset regarding intelligence. There are two main types of mindsets: a fixed mindset, which assumes that intelligence is unchangeable, and a growth mindset, which views intelligence as flexible and improvable through effort (Dweck, 1999). Lou and Noels (2019b) introduced the language mindset meaning system, which explains how different language mindsets relate to key motivators such as effort beliefs, achievement goals, fear of failure, and language anxiety. Evidence is mounting that adopting a growth mindset fosters greater achievements in academic settings, as learners with this perspective tend to be more motivated and resilient in their studies (Gouëdard, 2021). The recognition of a growth mindset as a key ingredient for success in language learning is resonating throughout academic circles, which is why an impressive rise has been observed in the breadth of research being conducted in this area within second language acquisition (SLA). Studies indicate that teachers have the ability to nurture a growth-oriented mindset in students. This type of outlook can assist learners in dealing with challenges and difficulties (Papi et al., 2019). Developing growth mindsets has the potential to improve learners' self-perception and emotional well-being (Zarrinabadi et al., 2021).

Given the critical role that emotions play in L2 acquisition, understanding how task enjoyment, defined as the positive emotional state experienced while engaging in specific language tasks, and task anxiety, defined as the negative emotional response felt during task performance, influence learners' language mindset is essential. While previous research has explored language anxiety and enjoyment as independent emotional factors, little attention has been paid to their direct effects on language mindset, particularly in task-based learning contexts. This study aims to fill this gap by examining how task-specific anxiety and enjoyment interventions impact Iranian EFL learners' views on language intelligence and L2 learning during a listening course. By doing so, the study provides insights into how emotional experiences shape learners' perceptions and attitudes, adding depth to our comprehension of the diverse individual factors influencing L2 acquisition.

2. Literature review

Emotions and cognitive factors play a crucial role in SLA. Research has explored the impact of FLE, FLCA, and language mindset on learners' motivation and performance. This section reviews key studies on these factors and their significance in SLA.

2.1. Research on foreign language enjoyment, and foreign language classroom anxiety

Psychologists have been studying how students' emotions impact general education for some time. Language anxiety emerged as the initial emotional factor to be recognized as a significant personal distinction among learners in the field of SLA research (Dewaele, 2022). Horwitz et al. (1986) put a spotlight on FLCA, presenting it as a distinct psychological concept separate from

general anxiety, specifically experienced in the context of language learning. Moreover, the development of their reliable instrument to measure FLCA triggered a period of extensive research within the field.

Research suggests a negative correlation between language anxiety and academic achievement. A meta-analysis by Teimouri et al. (2019) analyzed how anxiety relates to L2 achievement by drawing on a wide range of studies. They examined 97 studies across 23 countries and found an overall correlation of $r = -0.36$, indicating that higher anxiety levels negatively impact L2 learning, with factors such as assessment type and educational level influencing this relationship. Similarly, Zhang (2019) analyzed 46 studies, and a correlation of $r = -0.34$ was reported, indicating a negative interplay between anxiety levels and language performance; However, there was variation in the connection between anxiety and performance across different skill domains. Specifically, the link between anxiety and performance was the strongest in the context of listening ($r = -0.53$), while the association between anxiety and performance was less pronounced in reading and testing, with correlation values of $r = -0.23$ and $r = -0.27$, respectively. Similarly, Botes et al. (2020a) conducted a third meta-analysis of 59 studies, confirming a correlation of −0.39 between language anxiety and academic achievement, with varying effects among different language skills, particularly in listening and writing. Listening seems to be the most difficult among the four skills due to its complexity and fleeting nature. Unlike written text that can be revisited, sound fades away and cannot be easily reviewed, leaving the listener with little ability to influence or modify the input (Kim, 2000).

The rise of positive psychology at the beginning of the 21st century provided a fresh outlook within the discipline (Seligman & Csikszentmihalyi, 2000) and sparked a surge in SLA investigations into the impact of different emotions (MacIntyre & Mercer, 2014). In addition to the focus on anxiety as a negative emotion, the role of FLE as a motivating and uplifting emotion in language acquisition has sparked a growing body of research (Dewaele, 2022). Dewaele and MacIntyre (2014) revealed a meaningful moderate negative correlation ($r = -0.36$) that connects FLCA and FLE. This implies that there is a certain level of interrelationship between these two factors; However, it's worth noting that language anxiety doesn't always act as the opposite of language enjoyment, since high levels of anxiety can exist alongside high levels of enjoyment (Dewaele et al., 2018). Botes et al. (2022) conducted a meta-analysis analyzing 96 effect sizes from 28,166 participants and found a negative relationship between FLE and FLA. Additionally, FLE showed positive correlations with academic achievement and self-perceived success. Research has also linked FLE to emotional intelligence (Li, 2020; Li & Xu, 2019), and grit (Lee, 2022; Wei et al., 2019) as well as curiosity (Mahmoodzadeh & Khajavy, 2019). Since FLE is a fairly new factor, its connections with other variables are still developing, but initial results are encouraging for researchers exploring language learning through a positive psychology lens.

In the realm of TBLT, the impact of designing and implementing tasks in shaping learners' emotions has been a key area of study. Students have indicated that they experienced greater enjoyment alongside reduced anxiety when given more freedom in their tasks (Nakamura et al., 2021; Phung et al., 2021). Additionally, Lambert and Zhang (2019) found that students had lower anxiety levels when working on tasks that involved content generated by the learners themselves rather than by the teacher. Moreover, higher levels of enthusiasm, confidence, and interest were noted when participants had prior exposure to the task's subject matter and underlying concepts (Aubrey et al., 2022; Phung, 2017). Therefore, instead of thinking that emotions are just a stable trait that pops up because of a few specific factors, we should recognize that they actually come from a bunch of different factors all happening at the same time (Oxford & Gkonou, 2021). To develop a comprehensive understanding of the dynamics of various factors, it could be beneficial to take into account how the emotions of learners evolve over different time spans, during various tasks, and within individual task performances when planning a study on TBLT, and emotions (MacIntyre & Mercer, 2014; MacIntyre et al., 2019)

2.2. Research on language mindset

Several studies have suggested that mindsets may vary depending on the domain of application (Dweck, 2006; Dweck et al., 1995), and this idea has gained interest in recent studies on SLA (Khajavy et al., 2022; Waller & Papi, 2017). Some research revealed that adopting a growth mindset correlates positively with learning goals, effort beliefs, and mastery-oriented strategies. On the other hand, it shows an inverse relationship with anxiety, goals centered on preventing failure, and avoidance-based coping methods (Cutumisu & Lou, 2020; Dweck & Yeager, 2019; Lou & Zarrinabadi, 2022). Because of these findings and the potential impacts that a language mindset can have on SLA and learning in the classroom, it is a psychological concept that merits further exploration and potential action or application within an educational setting. The concept of a language mindset seems to be worth examining in more depth and addressing through educational interventions, given its implications for how students learn a new language.

So far, with the notion that people's mindsets can be altered through intervention (Blackwell et al., 2007; Wilson & English, 2017), only few research studies have examined how different interventions impact Language mindset within the SLA field. In a study conducted by Lou and Noels (2016), 150 university students participated in a randomized controlled trial. The results showed that after reading a mock article on growth mindset, students had a more constructive response to challenges, putting in more effort and showing a higher and increased motivation to persist in learning their L2. Similarly, Lou and Noels (2019a) reported that intervention on mindsets had a notable impact on intergroup anxiety, language-related rejection, along with expectations for adapting to a new culture. Moreover, Lou and Noels (2020), employing a similar research design (reading articles on language mindsets in a single session) as Lou and Noels (2016; 2019a), reported that implementing a growth-mindset intervention led to a decrease in the perceived rejection and avoidance of future interactions in ESL learners with limited English proficiency.

Molway and Mutton (2020) conducted a study in the UK where they presented students with research findings on mindsets. The researchers found that this intervention helped to promote growth mindsets and influenced how students reacted to challenges. In a longitudinal study, Lee et al. (2023) examined the effects of an English for Academic Purposes (EAP) writing program on the

language mindsets and academic writing skills of undergraduate students. Findings reveal significant improvements in students' growth language mindsets, particularly in general language intelligence beliefs, alongside notable enhancements in writing proficiency. Additionally, the results of the latent transition analysis suggested that the mindset intervention would lead to dynamic shifts in profile membership. In a more recent study, Bigverdi and Sabet (2024) examined the effects of online peer feedback (OPF) and online teacher feedback (OTF) on the writing development and language mindset of Iranian EFL learners. Using a quasi-experimental design, the researchers assigned 72 participants to three groups: OPF, OTF, and a control group. The findings revealed that while both feedback types improved writing performance, OPF positively influenced learners' language mindset, promoting a growth-oriented attitude toward language learning.

Research on language mindset interventions suggests that they can temporarily lead to brief changes in learners' mindset and motivation. However, in TBLT, insufficient focus has been given to the learner's role and individual differences, particularly in how emotions fluctuate during tasks (Almukhaild & King, 2023). While most studies have explored the effects of emotions over extended periods, task-specific emotions have received limited attention, especially in listening tasks, which research has shown to be highly anxiety-inducing (Botes et al., 2020a; Zhang, 2019). Given that listening requires real-time processing without the ability to revisit the content, it poses unique cognitive and emotional challenges, making it a suitable context for investigating task-related enjoyment and anxiety. Understanding how these emotions influence language mindset can help teachers and curriculum developers create more supportive learning environments, improving student engagement and motivation. By examining task enjoyment and anxiety interventions in a listening course, this study aims to provide insights into their impact on language mindset, contributing to more effective pedagogical practices and learner support strategies.

2.3. Research questions

This study tried to investigate the effects of task-specific enjoyment (the enjoyment that students experience performing a particular task) intervention and task-specific anxiety (the anxiety that students experience performing a particular task) intervention on the mindset of the students in a listening course. In summary, the research questions are as follows:

1. Does task-specific anxiety reduce listening anxiety among EFL learners?
2. Does task-specific enjoyment enhance listening enjoyment among EFL learners?
3. Does task-specific anxiety influence the language mindset among EFL learners in a listening course?
4. Does task-specific enjoyment influence the language mindset among EFL learners in a listening course?

3. Methodology

3.1. Participants

The study involved 75 male students, aged 16 to 18 (M= 16.9, SD= 0.53), who were organized into three intact classes. Participants were selected using convenience sampling, as the researchers had access to students from three schools located in Northwestern Iran. The students were similar in terms of intelligence, and they were all enrolled in schools for gifted students and came from comparable socioeconomic backgrounds. To evaluate the impact of the interventions, participants were randomly assigned to two experimental groups and one control group. All participants were Persian and Turkish bilinguals and they had passed English courses at high school. Based on their scores throughout the semester, they were all classified at an intermediate level of proficiency in English.

3.2. Research design

The study adopted a quantitative research design, including a quasi-experimental approach that utilized pre-test and post-test measures with control and experimental groups. This design allowed for the comparison of task-specific enjoyment and anxiety, as well as changes in learners' language mindset, between the groups exposed to different interventions within a listening course. The data collection consists of administering a set of standardized questionnaires to a sample of Iranian EFL learners.

3.3. Instruments

3.3.1. Foreign language enjoyment scale

Dewaele and MacIntyre (2014) developed the Foreign Language Enjoyment Scale (FLES) which is composed of 21 statements that are worded positively. The participants were required to provide feedback on these statements using a standard 5-point Likert scale. The FLES is structured in a way that groups the items into three categories, which reflect the various aspects of FLE. These categories comprise items that reflect the private dimension of FLE (learning experience), the social dimension (peers), and those that relate to the foreign language teacher. The FLE scale in its Persian adaptation (Elahi Shirvan et al., 2024) was implemented in this research. A pilot study with 18 students was performed and the measurement scale displayed strong internal consistency ($\alpha = 0.86$).

3.3.2. Foreign language classroom anxiety scale

Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz et al., 1986) is a widely used measure of anxiety experienced by learners in the classroom. The scale consists of 33 items and the respondents rate the degree to which they experience each item on a 5-point Likert scale, ranging from "strongly disagree" to "strongly agree". The Persian version of the FLCAS (Amiri & Ghonsooly, 2015) was used for this study. A pilot study with 18 students was undertaken, which showed that the measurement scale possessed significant internal consistency ($\alpha = 0.88$).

3.3.3. Language Mindset Inventory

Language Mindset Inventory (LMI) developed by Lou and Noels (2017) consists of 18 questions on a 6-point Likert scale. The LMI is organized into three distinct categories that represent different facets of language mindset. These categories consist of items that indicate beliefs about general language intelligence (GLB), beliefs about second language learning (L2B), and beliefs about age sensitivity and language learning (ASB). For this study, the Persian version of LMI (Khajavy et al., 2021) was employed. The measurement scale showed strong internal reliability ($\alpha = 0.83$) in a pilot study involving 18 students.

3.4. Procedure

First of all, to collect the data, ethics approval was obtained from the schools to conduct this study. The participants were randomly assigned to one of three groups: Control, Task Enjoyment (TE) Group, and Task Anxiety (TA) Group. The TA Group participated in a high-stakes listening task designed to simulate a real testing environment. This task was officially part of their mid-term assessment, contributing 4 out of 20 points toward their exam score. The task consisted of three challenging listening comprehension podcasts, sourced from the Vision Book series— the standard English textbook used in Iranian high schools (See Appendix A). After each podcast, students needed to address comprehension questions directly tied to the content. The podcasts included academic topics and formal dialogues at a level above the learners' current proficiency, which increased task complexity. Participants were explicitly informed of the importance of this task for their final grade, reinforcing the evaluative pressure. This intervention was aimed at inducing task-specific anxiety, in line with prior studies showing that listening tasks, particularly those tied to assessments, tend to elevate anxiety (Botes et al., 2020a, Zhang et al., 2019).

Alternatively, the TE Group engaged in a low-stakes listening task designed to foster enjoyment. These students were given the autonomy to choose from a list of familiar and personally relevant topics based on a needs analysis conducted prior to the study. Topics included everyday conversations, hobbies, travel stories, and humorous podcasts, all set at a lower difficulty level to match the TA group's task length and structure while lowering the cognitive load. Most students opted for a stand-up comedy podcast, followed by a series of light comprehension and reflective questions (See Appendix B). Importantly, students were informed that this activity was not graded and had no impact on their course assessment. This combination of topic familiarity, choice, and absence of grading was intended to create a more enjoyable and relaxed learning experience, helping to isolate the emotional response of enjoyment without the confounding influence of task difficulty.

Data collection occurred in December 2024 over two weeks. In the first week, LMI, FLES, and FLCAS were distributed among all participants. In the second week, TA and TE groups received the intervention and completed the LMI, FLCAS, and FLES immediately after receiving the intervention. It's important to point out that the Persian versions of the scales were checked by some experts and were used for collecting the data.

3.5. Data analysis

The data were analyzed using SPSS 26. Descriptive statistics were computed for all variables. To examine the effects of the interventions, a series of t-tests and a one-way ANOVA were conducted to compare post-test scores across the three groups (TA, TE, and Control) on task-specific anxiety, enjoyment, and the dimensions of language mindset (GLB, L2B, and ASB). Post-hoc comparisons using the Tukey HSD test were performed. The significance level was set at $p < .05$.

4. Results

Prior to running the t-tests and ANOVA, assumptions of normality and homogeneity of variances were tested. Shapiro-Wilk tests indicated that the data for all variables were normally distributed ($p > .05$). Levene's Test for Equality of Variances was conducted for each dependent variable, and results showed no significant violations of homogeneity, confirming that the assumption of equal variances was met for the analyses ($p > .05$).

Table 1 presents the descriptive statistics for task-specific enjoyment, task-specific anxiety, and the three dimensions of language mindset across the TE, TA, and Control groups.

Table 1. Descriptive statistics

	N	Mean	Std. Deviation	Std. Error

FLCA (pre-test)	TA Group	25	88.84	16.375	3.275
	TE Group	25	89.08	14.611	2.922
	C Group	25	88.92	19.687	3.937
	Total	75	88.95	16.790	1.939
FLCA (Post-test)	TA Group	25	99.24	16.205	3.241
	TE Group	25	85.24	13.128	2.626
	C Group	25	88.12	17.582	3.516
	Total	75	90.87	16.681	1.926
FLE (pre-test)	TA Group	25	64.72	12.821	2.564
	TE Group	25	64.92	10.924	2.185
	C Group	25	64.44	12.793	2.559
	Total	75	64.69	12.047	1.391
FLE (Post-test)	TA Group	25	61.64	12.446	2.489
	TE Group	25	73.04	13.834	2.767
	C Group	25	63.84	12.844	2.569
	Total	75	66.17	13.803	1.594
GLB (pre-test)	TA Group	25	28.44	3.190	.638
	TE Group	25	28.24	3.333	.667
	C Group	25	28.12	3.468	.694
	Total	75	28.27	3.289	.380
GLB (post-test)	TA Group	25	26.00	3.304	.661
	TE Group	25	31.12	2.651	.530
	C Group	25	28.16	3.287	.657
	Total	75	28.43	3.713	.429
L2B (pre-test)	TA Group	25	28.08	3.081	.616
	TE Group	25	28.12	3.257	.651
	C Group	25	27.80	3.000	.600
	Total	75	28.00	3.076	.355
L2B (post-test)	TA Group	25	25.16	3.223	.645
	TE Group	25	30.72	2.747	.549
	C Group	25	28.24	2.919	.584
	Total	75	28.04	3.718	.429
ASB (pre-test)	TA Group	25	28.32	3.579	.716
	TE Group	25	28.16	3.287	.657
	C Group	25	27.84	3.105	.621
	Total	75	28.11	3.290	.380
ASB (post-test)	TA Group	25	27.60	3.055	.611
	TE Group	25	28.84	2.154	.431
	C Group	25	27.52	3.280	.656
	Total	75	27.99	2.897	.334

4.1. Task-specific anxiety

To assess the impact of the intervention on task-specific anxiety levels, an independent-sample t-test was performed to compare the TA Group, which was exposed to a high-stakes, anxiety-inducing listening task, with the Control Group, which received no intervention (See Table 2).

Table 2. Independent samples test (post-test task anxiety scores)

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig	t	df	Sig. (2-tailed)
FLCA	Equal variances assumed	.607	.440	2.325	48	.024
	Equal variances not assumed			2.325	47.685	.024
FLE	Equal variances assumed	.027	.870	-.615	48	.541
	Equal variances not assumed			-.615	47.953	.541

The results of the independent-sample t-test revealed a statistically significant difference in task-specific anxiety levels between the TA Group and the Control Group. Specifically, participants in the TA Group reported significantly higher anxiety

(M = 99.24, SD = 16.205) compared to the Control Group (M = 88.12, SD = 17.582), t(48) = 2.325, p= .024, indicating that the intervention successfully induced a heightened anxiety response relative to the levels observed in the Control group. However, the level of FLE between the TA Group and the Control Group was not significantly different (p= .541), suggesting that the anxiety-inducing conditions did not notably diminish enjoyment levels.

4.2. Task-specific enjoyment

To further examine the emotional impact of the FLE intervention, another independent-sample t-test was conducted to compare the levels of task-specific enjoyment between the TE Group, which engaged in a low-stakes, autonomy-supportive listening task, and the Control Group, which did not receive any specific intervention. The data are presented in Table 3.

Table 3. Independent samples test (post-test task enjoyment scores)

		Levene's Test for Equality of Variances		t-test for Equality of Means		
		F	Sig	t	df	Sig. (2-tailed)
FLCA	Equal variances assumed	4.686	.035	-.656	48	.515
	Equal variances not assumed			-.656	44.417	.515
FLE	Equal variances assumed	.000	.988	2.437	48	.019
	Equal variances not assumed			2.437	47.738	.019

The statistical analysis demonstrated that the TE Group, which participated in the low-pressure, learner-centered listening task, experienced significantly greater task-specific enjoyment (M = 73.04, SD = 13.834) compared to the Control Group (M = 63.84, SD = 12.844). This difference reached statistical significance, t(48) = 2.437, p = .019, suggesting that the incorporation of learner autonomy, topic familiarity, and the removal of evaluative pressure meaningfully and significantly enhanced participants' enjoyment. However, the comparison of FLCA levels between the TE Group and the Control Group showed no statistically significant difference (p= .515), indicating that the enjoyment-enhancing conditions did not significantly reduce anxiety levels.

4.3. Language mindset outcomes

Following the confirmation that the data met the necessary statistical assumptions of normality and homogeneity of variances, a one-way ANOVA was performed to investigate potential differences across the TE, TA, and Control groups on each dimension of the language mindset construct (See Table 4).

Table 4. One-way ANOVA (post-test language mindset scores)

		Sum of Squares	df	Mean Square	F	Sig.
GLB	Between Groups	330.347	2	165.173	17.235	.000
	Total	1020.347	74			
L2B	Between Groups	387.920	2	193.960	21.994	.000
	Total	1022.880	74			
ASB	Between Groups	27.387	2	13.693	1.661	.197
	Total	620.987	74			

Considering the first subscale of LMI (i.e. GLB), a significant group difference was found for beliefs about general language intelligence between groups, F(2, 72) = 17.235, p = .000. Similarly, for L2 learning beliefs (L2B), there was a significant difference among the groups, F(2, 72) = 21.994, p = .000. However, the ANOVA analysis failed to find a significant difference in beliefs about age sensitivity (ASB) between the groups, F(2, 72) = 1.661, p = .197.

A Tukey post-hoc test was performed to determine which groups exhibited differences and to provide a more detailed insight into the variations among them. Table 5 outlines the findings, emphasizing the group pairs with statistically significant differences in GLB and L2B.

Table 5. Multiple comparisons of GLB and L2B (post-hoc test)

Dependent Variable	(I) Groups	(J) Groups	Mean Difference (I-J)	Std. Error	Sig.
GLB	TA Group	TE Group	-5.120*	.876	.000
		C Group	-2.160*	.876	.042
	TE Group	TA Group	5.120*	.876	.000
		C Group	2.960*	.876	.003
	C Group	TA Group	2.160*	.876	.042
		TE Group	-2.960*	.876	.003
L2B	TA Group	TE Group	-5.560*	.840	.000
		C Group	-3.080*	.840	.001
	TE Group	TA Group	5.560*	.840	.000
		C Group	2.480*	.840	.012
	C Group	TA Group	3.080*	.840	.001
		TE Group	-2.480*	.840	.012

* The mean difference is significant at the 0.05 level.

The post-hoc analysis revealed that the TE Group (M = 31.12) scored significantly higher than the TA Group (M = 26.00, p = .000) and significantly higher than the Control Group (M = 28.43, p = .003) in the GLB. On the other hand, the TA Group scored significantly lower than the Control Group (mean difference = -2.160, p = .042).

Considering L2B, the TE Group (M = 30.72) significantly outperformed the TA Group (M = 25.16, p = .000) and the Control Group (M = 28.24, p = .012). Furthermore, the post hoc test indicated that the TA Group scored significantly lower than the Control Group (mean difference = -5.560, p = .001).

5. Discussion

The current study sought to explore how task-specific enjoyment and task-specific anxiety interventions influenced the language mindset of Iranian EFL learners in a listening course. The findings revealed significant and nuanced insights, extending the existing literature in several meaningful ways. While the effects of long-term emotions such as trait anxiety and enjoyment have been well-documented (Botes et al., 2020b; Dewaele & MacIntyre, 2014), this study contributes to the underexplored area of task-specific emotional dynamics. Unlike prior research that has focused predominantly on classroom-level emotions across longer timeframes (MacIntyre et al., 2019; Oxford & Gkonou, 2021), this study emphasizes how discrete task-based emotional episodes can shape cognitive beliefs.

Concerning the first and the second research questions of the study, the interventions successfully manipulated task-specific anxiety and enjoyment. Specifically, the results highlighted that those in the TA Group exhibited significantly higher anxiety levels compared to the Control Group, while the TE Group reported significantly higher enjoyment than the Control Group. These findings are consistent with previous research by Botes et al. (2020a), who emphasized that listening tasks, because of their inherent cognitive demands and transient characteristics, often induce heightened anxiety among language learners. On the other hand, studies by Nakamura et al. (2021) and Phung et al. (2021) suggested that learner autonomy and task familiarity increase enjoyment—both of which were deliberately incorporated into the TE group's intervention design.

Consistent with prior research, the present findings reinforce the claims articulated by Dewaele and MacIntyre (2014) and Botes et al. (2022) that anxiety and enjoyment are not polar opposites but can co-exist independently. The present findings indicate that enjoyment-enhancing interventions, such as low-stakes tasks featuring humor and learner choice, can significantly bolster positive emotional responses even in an otherwise anxiety-prone domain like listening.

Concerning the third and the fourth research questions of the study, the results revealed that task-specific emotions were found to significantly affect two dimensions of learners' language mindset: GLB and L2B. Specifically, the TE Group demonstrated a substantial improvement in both GLB and L2B compared to the TA and Control groups. In contrast, the TA Group's beliefs about general language intelligence and L2 learning significantly declined relative to the other groups. These findings align with the growing body of research linking positive emotions such as enjoyment with adaptive cognitive and motivational patterns (Dewaele et al., 2023; MacIntyre & Mercer, 2014). More specifically, the results echo Lou and Noels (2016) and Bigverdi & Sabet (2024), who found that interventions promoting positive affect—whether through mindset instruction or peer feedback—can foster a growth-oriented attitude in learners.

Moreover, the negative impact of anxiety on learners' mindsets is consistent with earlier research (Cutumisu & Lou, 2020; Dweck & Yeager, 2019), which suggests that high anxiety correlates with maladaptive coping strategies and fixed mindsets. The finding that anxiety negatively and significantly influenced both GLB and L2B reinforces the notion that when anxiety is task-induced, particularly in evaluative contexts (as in the TA group), it can foster a fixed view of one's language-learning potential. However, neither intervention yielded a significant impact on ASB. This finding suggests that while task anxiety and task enjoyment interventions may influence other aspects of language mindset, they did not notably alter participants' perspectives on age sensitivity in the context of language acquisition. The lack of significant effects on ASB in this study could

indicate that participants' attitudes toward age-related sensitivity in language learning are deeply ingrained and may not easily change through short-term interventions targeting task anxiety and task enjoyment.

Additionally, while Lou and Noels (2016) demonstrated that growth mindset interventions can foster positive responses to general language challenges, the current study uniquely shows that task-induced emotional states—specifically enjoyment and anxiety in a single listening task—can similarly shape learners' mindset profiles. This situational perspective, grounded in CDST (MacIntyre et al., 2014), underscores the importance of understanding emotions as fluctuating, context-dependent variables with immediate cognitive repercussions.

Another noteworthy aspect of this study is its focus on listening tasks. While prior studies (Zhang, 2019; Kim, 2000) have identified listening as the skill most susceptible to anxiety due to its fleeting input and limited learner control, this research demonstrates that task design (i.e., stakes, topic familiarity, and autonomy) can substantially modulate learners' emotional experiences and their subsequent mindsets.

Theoretically, this study supports the notion that emotions in SLA should be framed within the CDST lens, given their rapid fluctuation and complex interaction with cognitive variables like mindset (Elahi Shirvan et al., 2020). Pedagogically, the results suggest that TBLT practitioners must attend not only to task complexity and learner proficiency but also to the emotional affordances of tasks. Specifically, fostering enjoyment through autonomy, relevant content, and non-evaluative contexts can cultivate more adaptive views on the process of language learning. Conversely, this study also highlights the risk of high-stakes, anxiety-inducing tasks. Even when used for assessment purposes, such tasks may reinforce fixed beliefs about language intelligence and undermine motivation—findings consistent with Ellis's (2003) argument that anxiety, when excessive, hampers learning.

6. Conclusion

This study investigated the effects of task-specific enjoyment and anxiety interventions on the language mindset of Iranian EFL learners within a listening course, providing novel insights into the immediate impact of task-induced emotions on learners' views on language learning. The findings revealed that fostering task-specific enjoyment through autonomy, familiar content, and low-stakes environments significantly promoted growth-oriented beliefs about general language intelligence and second language learning. Conversely, inducing task-specific anxiety with the use of high-stakes listening tasks resulted in more fixed and maladaptive language mindsets. These results highlight the critical role of task design in shaping learners' emotional experiences and subsequent cognitive perceptions within the TBLT framework.

The study reinforces the importance of integrating positive psychology principles into language education, particularly when designing listening tasks that are often prone to triggering anxiety. From a pedagogical standpoint, educators are encouraged to create emotionally supportive learning environments that prioritize learner autonomy and task relevance to cultivate more adaptive mindsets and enhance engagement. Theoretically, these findings contribute to the growing body of research advocating for the application of CDST in SLA, emphasizing the fluid and context-sensitive nature of learner emotions and language mindset.

While the study provides meaningful contributions, some limitations should be mentioned. One limitation of this study is its reliance on a male-only sample from a specific geographic and cultural context (Northwestern Iran). Future studies should explore more diverse populations, including female learners and different age groups, to ensure broader generalizability. Additionally, while this study focused on immediate post-task emotional and cognitive shifts, longitudinal designs could investigate the durability of these mindset changes over time. Further, future research could integrate qualitative data (e.g., learner diaries or interviews) to gain richer insights into the subjective emotional and cognitive experiences of learners during task performance.

7. Acknowledgments

We would like to express our heartfelt gratitude to all the participants who contributed to this study. Your involvement was invaluable. We also extend our thanks to the anonymous reviewers for their insightful feedback and constructive suggestions, which greatly enhanced the quality of this article.

8. References

Almukhaild, H., & King, J. (2023). Emotions in TBLT. In C. Lambert, S. Aubrey, & G. Bui (Eds.), *The role of the learner in task-based language teaching* (pp. 58-73). Routledge.

Amiri, M., & Ghonsooly, B. (2015). The relationship between English learning anxiety and the students' achievement on examinations. *Journal of Language Teaching and Research*, *6*(4), 855-865. https://doi.org/10.17507/jltr.0604.20

Aubrey, S., King, J., & Almukhaild, H. (2022). Language learner engagement during speaking tasks: A longitudinal study. *RELC Journal*, *53*(3), 519-533. https://doi.org/10.1177/0033688220945418

Bigverdi, A., & Sabet, M. K. (2024). The effects of online teacher feedback and online peer feedback on writing development and language mindset of the EFL learners. *Iranian Journal of English for Academic Purposes*, *13*(3), 1-17.

Blackwell, L. S., Trzesniewski, K. H., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, *78*(1), 246-263. https://doi.org/https://doi.org/10.1111/j.1467-8624.2007.00995.x

Botes, E., Dewaele, J.-M., & Greiff, S. (2020a). The foreign language classroom anxiety scale and academic achievement: An overview of the prevailing literature and a meta-analysis. *The Journal for the Psychology of Language Learning*, *2*(1), 26-56.

Botes, E., Dewaele, J.-M., & Greiff, S. (2022). Taking stock: A meta-analysis of the effects of foreign language enjoyment. *Studies in Second Language Learning and Teaching*, *12*(2), 205-232.

Botes, E. L., Dewaele, J.-M., & Greiff, S. (2020b). The power to improve: Effects of multilingualism and perceived proficiency on enjoyment and anxiety in foreign language learning. *European Journal of Applied Linguistics*, *8*(2), 279-306. https://doi.org/10.1515/eujal-2020-0003

Cutumisu, M., & Lou, N. M. (2020). The moderating effect of mindset on the relationship between university students' critical feedback-seeking and learning. *Computers in Human Behavior*, *112*, 106445. https://doi.org/10.1016/j.chb.2020.106445

Dewaele, J.-M. (2022). Enjoyment. In S. Li, P. Hiver, & M. Papi (Eds.), *The Routledge Handbook of second language acquisition and individual differences* (pp. 190-206). Routledge. https://doi.org/https://dx.doi.org/10.4324/9781003270546-16

Dewaele, J.-M., Botes, E., & Greiff, S. (2023). Sources and effects of foreign language enjoyment, anxiety, and boredom: A structural equation modeling approach. *Studies in Second Language Acquisition*, *45*(2), 461-479. https://doi.org/10.1017/S0272263122000328

Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, *4*(2), 237-274. https://doi.org/10.14746/ssllt.2014.4.2.5

Dewaele, J.-M., & MacIntyre, P. D. (2016). Foreign language enjoyment and foreign language classroom anxiety: The right and left feet of the language learner. In P. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215-236). Multilingual Matters. https://doi.org/10.21832/9781783095360-010

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language Teaching Research*, *22*(6), 676-697. https://doi.org/10.1177/1362168817692161

Dörnyei, Z. (2009). The L2 motivational self system. In Z. Dörnyei & E. Ushioda (Eds.), *Motivation, language identity and the L2 self* (pp. 9-42). Multilingual Matters. https://doi.org/doi.org/10.21832/9781847691293-003

Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality, and development*. Psychology Press.

Dweck, C. S. (2006). *Mindset: The new psychology of success*. Random House.

Dweck, C. S., Chiu, C.-y., & Hong, Y.-y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*(4), 267-285. https://doi.org/10.1207/s15327965pli0604_1

Dweck, C. S., & Yeager, D. S. (2019). Mindsets: A view from two eras. *Perspectives on Psychological Science*, *14*(3), 481-496.

Elahi Shirvan, M., Lou, N. M., & Taherian, T. (2021). Where do language mindsets come from? An ecological perspective on EFL students' mindsets about L2 writing. *Journal of Psycholinguistic Research*, *50*(5), 1065-1086. https://doi.org/10.1007/s10936-021-09787-y

Elahi Shirvan, M., Taherian, T., & Yazdanmehr, E. (2020). The dynamics of foreign language enjoyment: An ecological momentary assessment. *Frontiers in Psychology*, *11*, 1391. https://doi.org/10.3389/fpsyg.2020.01391

Elahi Shirvan, M., Taherian, T., & Yazdanmehr, E. (2024). Foreign language enjoyment: a longitudinal confirmatory factor analysis–curve of factors model. *Journal of Multilingual and Multicultural Development*, *45*(2), 224-242. https://doi.org/10.1080/01434632.2021.1874392

Ellis, A. (2003). Early Theories and Practices of Rational Emotive Behavior Therapy and How They Have Been Augmented and Revised During the Last Three Decades. In M. E. Bernard & W. Dryden (Eds.), *Advances in REBT: Theory, practice, research, measurement, prevention and promotion* (pp. 1-21). Springer International Publishing. https://doi.org/10.1007/978-3-319-93118-0_1

Ellis, R., Skehan, P., Li, S., Shintani, N., & Lambert, C. (2019). *Task-based language teaching: Theory and practice*. Cambridge University Press. https://doi.org/10.1017/9781108643689

Gouëdard, P. (2021). *Sky's the limit: Growth mindset, students, and schools in PISA. PISA 2018.* OECD Publishing.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*(2), 125-132. https://doi.org/10.2307/327317

Jackson, D. O. (2022). *Task-based language teaching*. Cambridge University Press.

Khajavy, G. H., MacIntyre, P. D., & Hariri, J. (2021). A closer look at grit and language mindset as predictors of foreign language achievement. *Studies in Second Language Acquisition*, *43*(2), 379-402. https://doi.org/10.1017/S0272263120000480

Khajavy, G. H., Pourtahmasb, F., & Li, C. (2022). Examining the domain-specificity of language mindset: a case of L2 reading comprehension. *Innovation in Language Learning and Teaching*, *16*(3), 208-220. https://doi.org/10.1080/17501229.2021.1956936

Kim, J.-h. (2000). *Foreign language listening anxiety: A study of Korean students learning English.* The University of Texas at Austin.

Lambert, C., & Zhang, G. (2019). Engagement in the use of English and Chinese as foreign languages: The role of learner-generated content in instructional task design. *The Modern Language Journal 103*(2), 391-411. https://doi.org/10.1111/modl.12560

Lee, H., Lee, J. H., & Scarcella, R. C. (2023). Influencing language mindsets and English writing competence through an EAP program: A longitudinal study with latent transition analysis. *Language Teaching Research*, *27*(1), 1-23. https://doi.org/10.1177/13621688221143258

Lee, J. S. (2022). The role of grit and classroom enjoyment in EFL learners' willingness to communicate. *Journal of Multilingual and Multicultural Development*, *43*(5), 452-468. https://doi.org/10.1080/01434632.2020.1746319

Li, C. (2020). A positive psychology perspective on Chinese EFL students' trait emotional intelligence, foreign language enjoyment and EFL learning achievement. *Journal of Multilingual and Multicultural Development*, *41*(3), 246-263. https://doi.org/10.1080/01434632.2019.1614187

Li, C., & Xu, J. (2019). Trait emotional intelligence and classroom emotions: A positive psychology investigation and intervention among Chinese EFL learners. *Frontiers in Psychology*, *10*, 2453. https://doi.org/10.3389/fpsyg.2019.02453

Lou, N. M., & Noels, K. A. (2016). Changing language mindsets: Implications for goal orientations and responses to failure in and outside the second language classroom. *Contemporary Educational Psychology*, *46*, 22-33. https://doi.org/10.1016/j.cedpsych.2016.03.004

Lou, N. M., & Noels, K. A. (2017). Measuring language mindsets and modeling their relations with goal orientations and emotional and behavioral responses in failure situations. *The Modern Language Journal 101*(1), 214-243. https://doi.org/10.1111/modl.12380

Lou, N. M., & Noels, K. A. (2019a). Promoting growth in foreign and second language education: A research agenda for mindsets in language learning and teaching. *System*, *86*, 102126. https://doi.org/10.1016/j.system.2019.102126

Lou, N. M., & Noels, K. A. (2019b). Language Mindsets, Meaning-Making, and Motivation. In M. Lamb, K. Csizér, A. Henry, & S. Ryan (Eds.), *The Palgrave handbook of motivation for language learning* (pp. 537-559). Springer International Publishing. https://doi.org/10.1007/978-3-030-28380-3_26

Lou, N. M., & Noels, K. A. (2020). Breaking the vicious cycle of language anxiety: Growth language mindsets improve lower-competence ESL students' intercultural interactions. *Contemporary Educational Psychology*, *61*, 101847. https://doi.org/10.1016/j.cedpsych.2020.101847

Lou, N. M., & Zarrinabadi, N. (2022). Mindsets. In Li, S., Hiver, P., & Papi, M. (Eds.), *The Routledge Handbook of second language acquisition and individual differences* (pp. 128-144). Routledge.

MacIntyre, P. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere* (pp. 24-45). McGraw-Hill.

MacIntyre, P., D., & Ayers-Glassey, S. (2021). Positive psychology. In T. Gregersen, & Mercer, S. (Ed.), *The Routledge handbook of the psychology of language learning and teaching (1st ed.)* (pp. 61-73). Routledge. https://doi.org/https://doi.org/10.4324/9780429321498

MacIntyre, P., & Gregersen, T. (2012). Affect: The Role of Language Anxiety and Other Emotions in Language Learning. In S. Mercer, S. Ryan, & M. Williams (Eds.), *Psychology for Language Learning: Insights from Research, Theory and Practice* (pp. 103-118). Palgrave Macmillan UK. https://doi.org/10.1057/9781137032829_8

MacIntyre, P. D., Dörnyei, Z., & Henry, A. (2014). Conclusion: Hot Enough to be Cool: The Promise of Dynamic Systems Research. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), *Motivational Dynamics in Language Learning* (pp. 419-429). Multilingual Matters. https://doi.org/doi:10.21832/9781783092574-025

MacIntyre, P. D., & Mercer, S. (2014). Introducing positive psychology to SLA. *Studies in Second Language Learning and Teaching*, *4*(2), 153-172.

MacIntyre, P. D., Ross, J., & Clément, R. (2019). Emotions are motivating. *The Palgrave Handbook of motivation for language learning*, 183-202. https://doi.org/10.1007/978-3-030-28380-3_9

Mahmoodzadeh, M., & Khajavy, G. H. (2019). Towards conceptualizing language learning curiosity in SLA: An empirical study. *Journal of Psycholinguistic Research*, *48*, 333-351. https://doi.org/10.1007/s10936-018-9606-3

Molway, L., & Mutton, T. (2020). Changing mindsets in the modern foreign languages classroom: An intervention combining intelligence theories and reading strategies. *The Language Learning Journal*, *48*(5), 598-612. https://doi.org/10.1080/09571736.2018.1554693

Nakamura, S., Phung, L., & Reinders, H. (2021). The effect of learner choice on L2 task engagement. *Studies in Second Language Acquisition*, *43*(2), 428-441. https://doi.org/10.1017/S027226312000042X

Oxford, R. & Gkonou, C. (2021). Working with the Complexity of Language Learners' Emotions and Emotion Regulation Strategies. In R. Sampson & R. Pinner (Ed.), *Complexity perspectives on researching language learner and teacher psychology* (pp. 52-67). Multilingual Matters. https://doi.org/10.21832/9781788923569-005

Papi, M., Rios, A., Pelt, H., & Ozdemir, E. (2019). Feedback-seeking behavior in language learning: Basic components and motivational antecedents. *The Modern Language Journal* *103*(1), 205-226. https://doi.org/https://doi.org/10.1111/modl.12538

Phung, L. (2017). Task preference, affective response, and engagement in L2 use in a US university context. *Language Teaching Research*, *21*(6), 751-766. https://doi.org/10.1177/1362168816683561

Phung, L., Nakamura, S., & Reinders, H. (2021). The effect of choice on affective engagement: Implications for task design. In H. Phil, H. A.-H. Ali, & M. Sarah (Eds.), *Student engagement in the language classroom* (pp. 163-181). Multilingual Matters. https://doi.org/doi:10.21832/9781788923613-012

Reeve, J. (2024). *Understanding motivation and emotion* (8 ed.). John Wiley & Sons.

Samuda, V. (2015). Tasks, design and the architecture of pedagogical spaces. In M. Bygate (Ed.), *Domains and directions in the development of TBLT: A decade of plenaries from the international conference* (pp. 271-302). John Benjamins Publishing Company. https://doi.org/10.1075/tblt.8.10sam

Seligman, M., & Csikszentmihalyi, M. (2000). Positive psychology: An introduction. *The American Psychologist*, *55*, 5-14. https://doi.org/10.1037/0003-066X.55.1.5

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*(2), 363-387. https://doi.org/10.1017/S0272263118000311

Waller, L., & Papi, M. (2017). Motivation and feedback: How implicit theories of intelligence predict L2 writers' motivation and feedback orientation. *Journal of Second Language Writing*, *35*, 54-65. https://doi.org/10.1016/j.jslw.2017.01.004

Wei, H., Gao, K., & Wang, W. (2019). Understanding the relationship between grit and foreign language performance among middle school students: The roles of foreign language enjoyment and classroom environment. *Frontiers in Psychology*, *10*, 1508. https://doi.org/10.3389/fpsyg.2019.01508

Wilson, A.E., English, J.A. (2017). The motivated fluidity of lay theories of change. In: Zedelius, C., Müller, B., Schooler, J. (eds) *The science of lay theories* (pp. 17-43). Springer. https://doi.org/10.1007/978-3-319-57306-9_2

Zarrinabadi, N., Rezazadeh, M., Karimi, M., & Lou, N. (2021). Why do growth mindsets make you feel better about learning and your selves? The mediating role of adaptability. *Innovation in Language Learning and Teaching*, *16*(3), 249–264. https://doi.org/10.1080/17501229.2021.1962888

Zhang, X. (2019). Foreign language anxiety and foreign language performance: A meta-analysis. *The Modern Language Journal*, *103*(4), 763-781. https://doi.org/10.1111/modl.12590

9. Appendices

9.1. Appendix A

Comprehension questions of TA group

Listening

Audio file 1: **Listen to the speaker and answer the questions.**

1. The woman did NOT want to buy any.....in the shop.
a) cheese b) rice c) milk d) tea

2. The woman was in a foreign country for ..... months.
a) four b) two c) three d) five

3. She walked around the shop for half an hour.
a) True b) False

Audio file 2: **Listen to the conversation and answer the questions.**

4. Based on the conversation, which one is correct?
a) Both of them will probably take part in the exam.
b) He doesn't have time to help his friend.
c) The exam is so easy that it doesn't need hard work.
d) Just one of them paid attention to the notice.

5) There will be an Olympiad Exam next ..... .
a) winter
b) spring
c) autumn
d) summer

Audio file 3: **Listen to the conversation and answer the questions.**

6. Which one does Marsha do on Sundays?
a) playing tennis b) sleeping late c) visiting friends d) swimming

7. Marsha goes to school on weekdays.
a) True b) False

8. Marsha exercises ..... days a week.
a) four b) three c) two d) five

9.2. Appendix B

Comprehension questions of TE group

Provide short answers to the following questions:

1. What was your favorite part of the podcast?
2. Did you find any jokes confusing? Which ones?
3. How did the podcast make you feel? Happy, relaxed, or something else?
4. Would you recommend this podcast to a friend? Why or why not?
5. What did you like most about the comedian's style?
6. Was there anything you didn't like about the podcast?
7. Did you learn anything new from the episode? What was it?

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 58-72

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.9119.1116

Developing and Validating an Attentional Literacy Model for Language Learners: A Fuzzy Delphi Study

Maryam Zarrabi[1]
Mojtaba Mohammadi[2*]
Zohreh Seifoori[3]
Agnieszka Palalas[4]

[1]Ph.D. Candidate in Applied Linguistics, Department of English, SR.C., Islamic Azad University, Tehran, Iran
[*2]Assistant Professor of Applied Linguistics, Department of English Language Teaching, WT.C., Islamic Azad University, Tehran, Iran
[3]Associate Professor of Applied Linguistics, Department of English, SR.C., Islamic Azad University, Tehran, Iran
[4]Associate Professor of Applied Linguistics, Department of Distance Education, Athabasca University, Athabasca, Canada

ARTICLE HISTORY

Received: 08 December 2024
Revised: 13 January 2025
Accepted: 24 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: mohammadi.mojtaba@wtiau.ac.ir

ABSTRACT

In a world with multifarious sources of disarrays, learners must be equipped with many literacies to handle their learning challenges. As a macroliteracy, attentional literacy refers to the learners' awareness of maintaining their attention on the information from the self, others, and the environment without being judgmental toward differing viewpoints and contexts. Assuming attention as not purely cognitive but social, cultural, affective, and technological, the researchers could not find any theoretical model underpinning different dimensions of attentional literacy in the literature. This study aims to develop a conceptual model of attentional literacy for English language learners. Using a Fuzzy Delphi Method, we surveyed the related literature to identify the factors of attentional literacy. We proposed an initial model with six factors and a list of measures which, after preliminary proofreading and piloting, was administered to a group of 20 experts in two rounds. The results of Fuzzy Delphi led to the conceptualization of a new attentional literacy model comprising two dimensions (internal and external), five factors (cognitive, socio-emotional, environmental, technological, and pedagogical), and 40 measures. The findings cast light on the theoretical foundations of attentional literacy and elucidate the dimensions, factors, and measures for language learners. Pedagogical contributions are discussed for researchers in language education.

KEYWORDS: Attentional literacy model; Fuzzy Delphi method; Cognitive factors; Socio-emotional factors; Twenty-first-century literacy

1. Introduction

The new millennium emerged with new complexities brought to the everyday life of the students which are the sources of many disarrays. The multiple sources of information have overwhelmed their personal and academic life which can easily hinder the learning process. The bulky literature delineated it to such factors as motivation (e.g., Dornyei, 2020; Dornyei & Ushida, 2021), willingness to communicate (e.g., Henry & MacIntyre, 2024; MacIntyre & Wang, 2021), learning style/strategies (e.g., Griffiths, 2018; Oxford, 2017), anxiety (e.g., Horwitz, 2010; Teimouri et al., 2019), etc. However, some studies have revealed that one factor which has remained overlooked and unaddressed is attention (e.g., Komorowska, 2021). It is further believed that inattention is among the major sources of students' deficiency in their academic performance (e.g., Arnold et al., 2020; Gray et al., 2017; Mohebbi, 2023). Some also affirmed that corrective feedback types embedded in negotiated interaction can help learners notice the erroneous language (Long, 1996, Ziaei & Dabaghi, 2023).

Cite this article: Zarrabi, M., Mohammadi, M., Seifoori, Z. and Palalas, A. (2025). Developing and validating an attentional literacy model for language learners: A Fuzzy Delphi study. *Applied Linguistics Inquiry*, *3*(1), 58-72. doi: 10.22077/ali.2025.9119.1116

The process of second language acquisition like other type of learning, entails factors which are both internal and external. Besides social, cultural, and political factors, there are some learner internal factors which deal with the processes enabling them to streamline their language learning. From a cognitive perspective in second language acquisition, the mental process of language learning entails "general approaches to learning adopted by the learners" and "specific … operations involved in different stages of acquisition" (p. 434) which Ellis (2008) referred to them as macro- and micro-processes respectively. Macro-processes like intentional/incidental learning or implicit/explicit learning have received considerable attention from scholars in SLA. However, the micro-processes like attention have not attracted the deserving attention of the studies. This might be due to the fact that the concept seems more of an interdisciplinary issue which is concerned with the fields like psychology, neurology, and physiology rather than solely pedagogy and education. Validating this model is essential as it enhances understanding of attentional dynamics—cognitive, social, and technological—impacting language learning process. By establishing a comprehensive framework, the study contributes to pedagogical strategies that foster effective learning environments in increasingly complex educational contexts. This study is also significant in that it primarily addresses attentional literacy as a new multidimensional concept through the prismatic lens of cognitive, psychological, ecological, pedagogical, and technological perspectives in language learning which has not been, to the humble knowledge of the researchers, probed yet. Moreover, following the recommendation of Sterling et al. (2023), this study adopted Delphi method which is limitedly used in applied linguistics and its "iterative process and community-driven nature of knowledge creation" (p. 9) can provide the field of applied linguistics "with additional validity arguments" (p. 9).

2. Literature review

There is no unanimous agreement among the scholars in education at large on the term attention. According to James (1890), attention is defined as "taking possession of the mind, in clear and vivid form, of one out of what seem several simultaneously possible objects or trains of thought" (pp. 403-404). Later scholars defined it as the degree a person orients or channels his senses to an object, event, or information and ignores other stimulants in the environment (Ott, 1994). American Psychiatric Association (2013) defined behavioral inattention as a number of observable behaviors which indicate distractibility, challenges in organization, forgetfulness, and difficulties in following directions and attending to pertinent stimuli. These symptoms either mild or severe are very crucial in the process of learning. In recent years, Melo et al. (2024) defined it as "a cognitive process that allows selective concentration on one aspect of the environment while ignoring others" (p. 835). As can be inferred from all the above definitions, the factors of concentration, selection, and a single object or action to attract attention are all common in these propositions.

Attention plays a crucial role in second language (L2) learning as it influences the processing and acquisition of new linguistic information. The ability to effectively allocate attentional resources can significantly impact language learning outcomes. Historically, the concept of attention was represented in terms of consciousness. Unlike the behaviorists who took a positivistic approach claiming no integration of conscious awareness in their learning framework, cognitivists view attention as a conscious process in which working memory is involved. Many related theories were proposed by the proponents of cognitivism such as acquisition versus learning (Krashen, 1981), noticing hypothesis (Schmidt, 1994), controlled versus divided attention (Eysenck, 2001), theory of attention (Tomlin & Villa, 1994), input processing (VanPatten, 1996), and multiple-resources model (Robinson, 2003).

Besides the intrinsic factors like the role of attentional control, working memory capacity, and executive functions in shaping attention, extrinsic factors can also play a key role in maintaining and managing attention in the process of language learning which underlies the tenet of the sociocultural paradigm shaping the theoretical foundation of this study. Sociocultural perspectives on attention focus on the socio-cognitive and external factors that influence attentional processes. These perspectives consider how attention is shaped by the social and cultural contexts in which individuals are embedded. Factors such as cultural norms, social expectations, and environmental influences are believed to impact attention to a considerable extent (Schepers, 2007). According to Piaget's theory of sociocultural theory, learners actively construct knowledge through assimilation and accommodation. Attention is considered a vital component in this process as it enables learners to select and focus on relevant information, facilitating the integration of new knowledge with existing cognitive structures (Chapman, 1987). In the sociocultural paradigm, attention is viewed within the context of social interaction and cultural mediation. Vygotsky's sociocultural theory underlines the role of interaction in the development of cognition. Attention, in this framework, is seen as a socially mediated process that is influenced by the cultural tools and practices of a particular community (Vygotsky, 1994). According to Vygotsky, individuals acquire attentional skills through their interactions with more knowledgeable others, who provide guidance and scaffolding to support their learning. Learners notice and attend to specific aspects of the learning environment based on the guidance and support provided by more knowledgeable others. Through social interaction, learners develop the ability to notice and attend to culturally relevant information, which contributes to their cognitive development (Rogoff, 1990).

A question naturally arises here: Is it only a matter of *having* attention or *managing* it? With the turn of the century and the essential role of literacies in fulfilling the aims of everyday life, attention is viewed as a form of awareness that might foster mindful living. According to Pegrum and Palalas (2021), attentional literacy (AL) is a term that encompasses the ability to effectively manage and direct one's attention in an increasingly distracting world. In contemporary society, individuals are incessantly exposed to a constant influx of information from diverse sources, such as social media, television, and advertisements. This persistent influx of stimuli can often lead to attentional overload, making it challenging for individuals to focus on important tasks and make informed decisions. Therefore, AL plays a crucial role in enabling individuals to navigate this information-rich environment and make optimal use of their attentional resources.

In recent years, AL has gained significant attention from researchers and scholars in various fields, including psychology, neuroscience, and education. Several studies have explored the impact of attention on cognitive processes, such as learning, memory, and decision-making (Berger et al., 2023; Español-Martín et al., 2023; Miyadera, 2024; Nurmuhammad, 2023; Stewart & Swanson, 2024; Stevens et al., 2024). These studies have indirectly highlighted the importance of AL in enhancing these cognitive functions and improving overall performance. AL is a multidimensional construct that encompasses the ability to effectively manage and direct one's attention. It plays a critical role in various aspects of life, including education, cognitive processes, emotional well-being, and mental health. The growing body of research on AL highlights its significance in today's information-rich society (e.g., Palalas, et al., 2024).

In a world flooded with information, especially through digital media and exposure to countless sources, attention deficits or distractions are more prevalent than ever. In education, AL is investigated to be the source of educational/academic failure (Pegrum & Palalas, 2021) and challenges such as depression and anxiety, addictive behaviors, and prosocial attitudes (Bonnardel et al., 2018; Melo et al, 2020) which need further attention. Moreover, in the cognitive era, attention, or what they called attentional control theory (Eysenck & Derakshan, 2011), was traditionally seen as a unidimensional concept focusing on the quality of information processing from sensory input to short-term memory and finally to higher-level information processing. This narrows the concept of attention into only one of its subsections defining it solely within the realm of cognition. Literature is rich investigating attentional control theory in language learning environment (e.g., Dong & Li, 2020; Finneran et al., 2009; Martínez-Vicente et al., 2023; Taghavi-Nejad et al., 2024). In recent years, under the influence of sociocultural paradigm, attention is not seen as a single unidimensional system, but as a complex idea made up of different parts, called attentional systems (Issa & Morgan-Short, 2019). These include an external system that is affected by outside cues related to how we sense things and an internal system that is shaped by the information we create ourselves. This model has been applied quite recently to investigate the effect of variations of input enhancement in L2 learners' vocabulary (Liu et al., 2024). In the sociocultural paradigm, attention is viewed as a multidimensional concept (Yip, 2023; Yip et al, 2023) which entails psychological, affective, cultural, and ideological perspectives as well. Meager literature is available on this last issue. Furthermore, with the recent rise of attentional literacy, as a newly developed concept, it appears to be an underexplored area within the field of second language acquisition. A comprehensive framework that conceptualizes the underlying dimensions and factors is also lacking, one that considers it not merely as a cognitive concept but also as an emotional, social, cultural, or technological one. This conceptualization is the first step to developing a scale to measure one's level of attentional literacy. To this end, the reported study primarily aimed to identify the dimensions, factors, and measures of attentional literacy by adopting a fuzzy Delphi method and subsequently providing a conceptual framework.

2.1. The initial proposed model

Based on the thorough review of the literature adopting a phylogenetic approach and presenting the concept of attention from a historical, theoretical, and developmental perspective as briefly reported in the previous section, the researchers have attempted to conceptualize a new model of attentional literacy. The proposed model consists of six main factors: cognitive, mindfulness-related, technological, psychological, environmental (social & physical), and pedagogical.

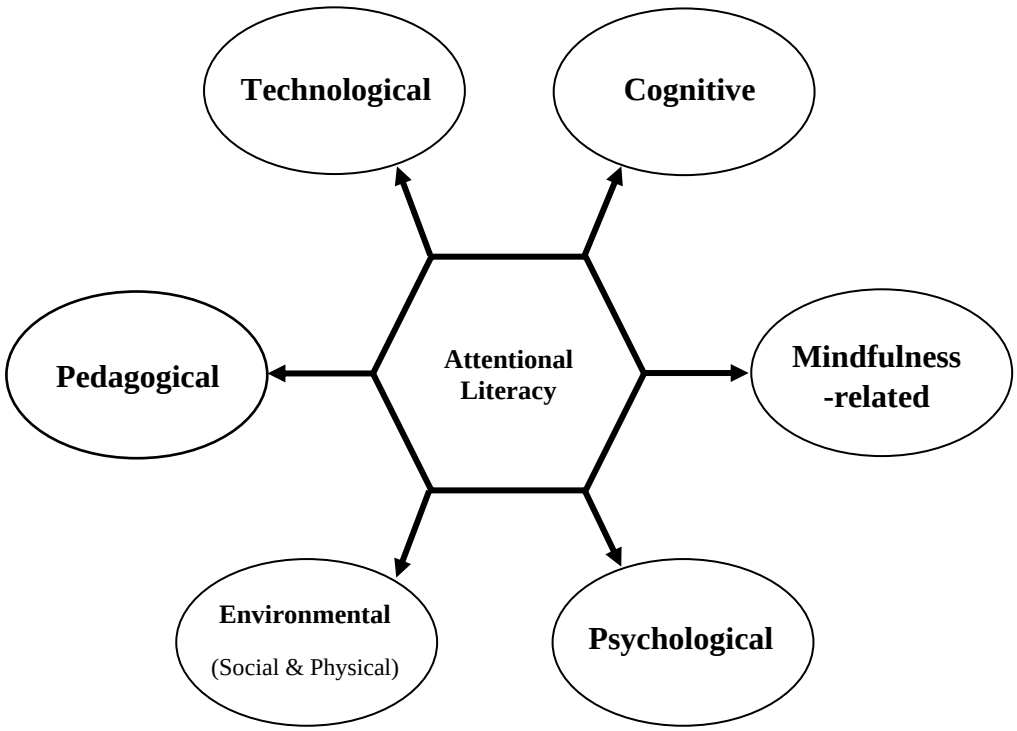

Figure 1. The initial attentional literacy conceptual model

2.2. Fuzzy Delphi method

The Fuzzy Delphi Method (FDM) is a research method which marries the Delphi method with fuzzy theory. The history of the Delphi Method dates back to the 1950s with RAND Corporation and was later presented by Dalkey and Helmer in 1963. In the Delphi Technique, the researcher seeks the consensus of a group of experts' opinions by systematic administration of a questionnaire to collect their judgment and justification. In many such situations, the experts' judgment cannot be understood in the form of definitive values or numbers. In other words, quantitative values and numbers fall short of accurately describing real-world systems due to the inherent ambiguity and uncertainty present in the judgments of those deciding within an expert panel. To overcome this problem, the theory of fuzzy sets or fuzzy theory presented by Zadeh (1965) can be used to deal with this ambiguity and uncertainty in the decision-making process. FDM is a kind of Delphi method that incorporates fuzzy theory to diminish the subjectivity of the results of the Delphi method by converting linguistic scale into fuzzy scale, item verification (fuzzification), and defuzzification. Like Delphi method which is used quite rarely in language education (e.g., Bulusan, 2024), the FDM has recently been adopted in studies with an exploratory nature within the field of language education (e.g., Ghani et al., 2024; Luo et al. 2024).

To fulfill the purpose of the study, the following questions were posed:

RQ1: What are the dimensions, factors, and measures of the attentional literacy conceptual model among language learners based on the expert agreement?

RQ2: What are their positions in the newly proposed model?

3. Methodology

3.1. Design

This is a multi-phase fuzzy Delphi study intended to validate the measures synthesized from the relevant literature. The study included two rounds of commenting from a panel of experts and analyzing the results by the researchers. It required two rounds to reach a final consensus about the dimensions, factors, and measures of the concept of attentional literacy. Attentional literacy was first raised in the conceptual research by Palalas and Pegrum (2021) who revised the term and reconceptualized the notion in 2021 and subsequently researched its application in online learning three years later (Palalas et al., 2024). Their definition of attentional literacy was a trigger to delve into the literature and survey its theoretical underpinnings along with the empirical research in the related fields.

3.2. Expert panel

The expert panel was comprised of 20 university professors teaching and researching in the sub-fields of applied linguistics, TESOL/TEFL, educational psychology, and distance education. Maximum variety in the panel members' expertise was observed to ensure the validity of the final model. They were selected based on the availability sampling method out of 35 scholars who were contacted and who opted in based on the letter of consent. Some did not reply despite their initial agreement. The inclusion criteria for the fuzzy Delphi panel members were extensive expertise in the field, being recognized in their areas of research interest, and over eight years of experience. The frequency of the panel members' professional expertise and areas of research is demonstrated in Table 1.

Table 1. List of experts involved in the Delphi technique and frequency

No.	Professional Expertise and Area of Interest	F
1	TEFL (Teaching English as a Foreign Language)/ TESOL	6
2	L2 Teacher Education	3
3	Pragmatics	1
4	Second Language Testing & Assessment	2
5	Discourse Analysis	2
6	Multiculturalism & Translation Studies	1
7	Technology-Assisted Learning / Distance Education	4
8	Psycholinguistics	1
9	Neurolinguistics	1
10	English for Academic/Specific Purposes	1
11	General/Educational Psychology	2
12	Sociolinguistics	1

There were 10 male and female panelists. They were both national and international scholars with the age between 39 to 74 within the age groups of 60-70 (%10), 50-60 (%55), 40-50 (%30), and 30-40 (%5). They were faculty members in 10 universities in Iran, Turkey, Canada, and Australia with the academic rank ranging from assistant professor (%30), associate

professor (%55), and professor (%15). They had an average of 23 years of teaching and research experience in their areas of research interest. Table 2 summarizes the demographic information of the panelists.

Table 2. Experts' demographic information

Demographic information	Details	Frequency	Percentage
Age	Below 40	1	5
	40-50	6	30
	51-60	11	55
	61-70	2	10
Gender	Male	10	50
	Female	10	50
Nationality	Iran	17	85
	Turkey	1	5
	Australia	1	5
	Canada	1	5
Academic rank	Assistant Professor	6	30
	Associate Professor	11	55
	Professor	3	15
Years of teaching/researching experience	Below 10	2	10
	10-20	6	30
	21-30	12	60

3.3. Procedure of data collection and analysis

The study entailed four major stages: literature review, expert assessment (round 1), expert assessment (round 2), and final analysis.

First, the researchers thoroughly and systematically reviewed the related literature until the point of saturation when there were no more theoretical frameworks or models available; hence, research in different disciplines was to be scrutinized to find appropriate dimensions and factors and develop the measures. Having completed the comprehensive review, six factors were identified and 64 measures were developed. The initial factors were cognitive, mindfulness-related, technological, psychological, environmental (social & physical), and pedagogical. To refine the factors and measures prior to the first fuzzy Delphi round, a pilot study was conducted. The factors and measures were consulted with four of the panel members, two of whom were the initiators of the concept of attentional literacy. After their review process, the issues were discussed through either online or face-to-face sessions with the researchers. The result of this pilot study was the revision and reclassification of the dimensions and factors into internal (cognitive, socio-emotional) and external (technological, pedagogical, environmental) factors and the deletion of 7 measures. The remaining five factors and 57 measures were converted to a digital questionnaire using Google Forms.

In the second stage, the refined questionnaire was then distributed to the members of the expert panel. They were given a two-week deadline which was then extended to four weeks to be able to accommodate their hectic schedules. The questionnaire consisted of three sections: demographic information survey, attentional literacy measures, and final comments. For the demographics, the panelists were asked to provide their full name, nationality, affiliation, age, years of experience, and areas of expertise. To raise the panelists' awareness of the attentional literacy concept, the AL measures section began with a brief background of the concept, alongside its theoretical definition of the concept of attentional literacy, and the direction to provide feedback on the measures. This section was added to the beginning of the questionnaire to raise the panelists' awareness regarding the concept and its background prior to reviewing and commenting on the measures. The panelists were supposed to give their assessment of the measures by ranking them based on a 7-point Likert scale (1= a strongly inappropriate measure, 2= an inappropriate measure, 3= a slightly inappropriate measure, 4= a slightly appropriate measure, 5= a moderately appropriate measure, 6= an appropriate measure, and 7= a strongly appropriate measure). After each measure, a comment box was included for the panelists' viewpoints and their response justification. Furthermore, at the end of the questionnaire, the experts were requested to add any additional missing factors and/or measures and their justification for them. After receiving the responses, the Microsoft Excel sheet was extracted and downloaded for analysis. Two sources of data were considered for the analysis: the written comments below each measure and in the final remarks section at the end of the questionnaire, and the panelists' responses to the quality of the measures based on a 7-point Likert scale. The comments below each measure (if there were any) were thoroughly read and discussed among the researchers. They varied from minor addition, deletion, or change of a word, and an item rewording, to complete inclusion or exclusion of an existing item, or adding a new item. Then, the researchers decided whether that comment was applicable or not. Peer review of at least two of the researchers was involved to guarantee the credibility of the data analysis. As for the responses of the panelists to the measures, a Fuzzy Delphi Method (FDM) procedure was adopted. It consisted of converting linguistic scale into fuzzy scale, item verification (fuzzification), and defuzzification which are reported in detail in the results section.

After conducting the FDM data analysis and item modifications, the newly revised questionnaire was sent to the expert panel for the second round of screening the items. The same procedure was conducted. The links to the modified questionnaire were sent via email to the members and a two-week deadline was announced. An Excel file including the aggregate results of all the panel members' responses to the items of the questionnaire in round 1 was also sent to the panelists to establish a consensus regarding the measures. When the responses were received, a similar data analysis procedure was conducted. The comments for the measures were discussed and decisions on any suggested modifications were made regarding the existing measures. Following that, the Excel sheet of the expert's opinion about the measures was downloaded and FDM analysis was done.

The fourth stage had no data collection. It mainly entailed the comparison of the results of FDM in the first and second rounds of collecting data from the experts. Based on the results of this comparison, the final factors and measures of the questionnaire were screened and verified. Figure 2 depicts the schematic representation of the procedures of data collection and analysis.

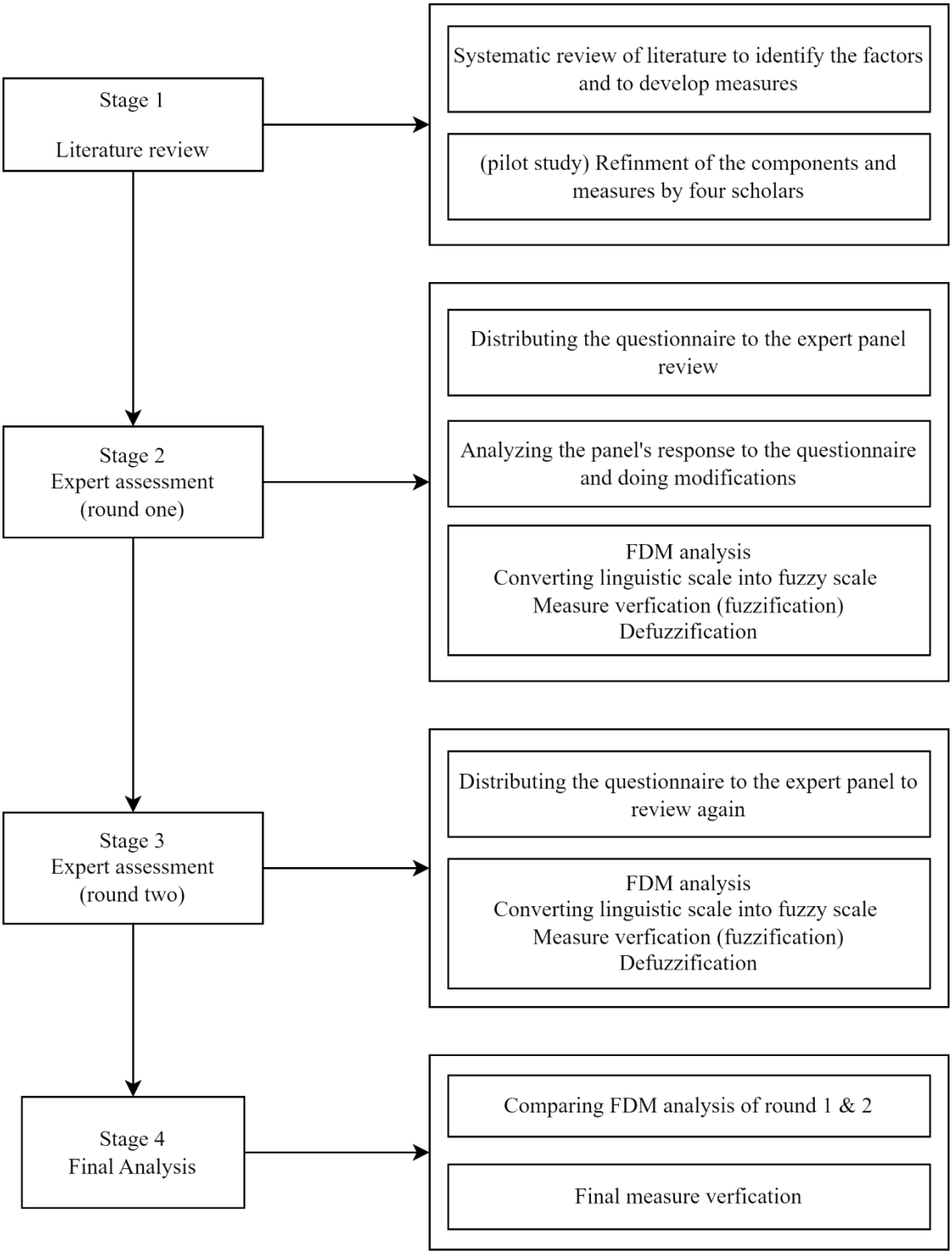

Figure 2. Data collection and analysis procedure

4. Results

In order to analyze the data in the FDM, the responses from the expert panelists were thoroughly analyzed adopting the steps suggested by Marlina et al. (2022) and Sensuse et al. (2018). These steps are used to analyze the FDM:

4.1. Converting the linguistic variables into fuzzy numbers

In this first phase, the Triangular Fuzzy Numbers (TFN), the combinations of three fuzzy numbers, were used to show the fuzziness of the expert opinion regarding the efficiency of the measures. Fuzzy numbers are the generalization of the human inexact use of real numbers in decision-making. It refers not to a sharp and crisp value but to a connected set of possible values, ranging between 0 to 1, which can help reach the utmost exactitude of the experts (Manakandan et al., 2017; Marlina et al., 2022). The expert opinions were translated into these three numbers. (Table 3)

Table 3. 7-point instrument scale

Instrument scale	Linguistic Variable	Triangular Fuzzy Numbers (m1, m2, m3)
7	Strongly Appropriate	(0.9, 1.0, 1.0)
6	Appropriate	(0.75, 0.9, 1.0)
5	Moderately Appropriate	(0.5, 0.75, 0.9)
4	Slightly Appropriate	(0.3, 0.5, 0.75)
3	Slightly Inappropriate	(0.1, 0.3, 0.5)
2	Inappropriate	(0.0, 0.1, 0.3)
1	Strongly Inappropriate	(0.0, 0.0, 0.1)

4.2. Items acceptability

In finding the acceptable items, the threshold value and the percentage of expert agreements were to be calculated. The former, threshold value, refers to the experts' acceptable level of agreement on a measure which was set to be < 0.2. A threshold value for all measures, following Al-Rikabi and Montazer (2024) was calculated by finding the difference between the fuzzy values average and the fuzzy numbers of the expert response based on the following formula:

$$\text{Threshold value } (m, n) = \sqrt{1/3\ [(m1 - n1)2 + (m2 - n2)2 + (m3 - n3)2]}$$

The latter, expert agreement, refers to the frequency of the experts' consensus. Any measure was accepted if the value was more than 75% (Chang et al., 2000); otherwise, the second round is needed to see if the measure is verified or not.

$$\text{expert agreement } (\%) = \frac{\text{the frequency of threshold values} \leq 0.2}{\text{Number of experts}}\ 100$$

4.3. Defuzzification

It includes converting the fuzzy values to a precise and crisp value considering the fuzzy set. In doing so, the following formula can be applied:

$$\text{Defuzzification value} = 1/3 * (m1 + m2 + m3)$$

Based on these steps, the data collected from 20 experts' responses to the measures of the attentional literacy scale in two rounds were analyzed. Table 4 summarizes the findings of the FDM analysis including triangular fuzzy numbers and experts' agreement of the first round.

Table 4. The results of round 1 on the proposed model

Measures	Triangular Fuzzy Numbers			Defuzzification Value	Expert agreement %	Measures	Triangular Fuzzy Numbers			Defuzzification Value	Expert agreement %
	m1	m2	m3				m1	m2	m3		
1	0.768	0.908	0.982	0.886	100	**30**	0.567	0.703	0.829	0.699	81
2	0.617	0.742	0.861	0.740	92	**31**	0.261	0.371	0.521	**0.384**	**56**
3	0.706	0.834	0.937	0.826	95	**32**	0.681	0.808	0.918	0.802	88
4	0.725	0.834	0.921	0.827	96	**33**	0.817	0.918	0.984	0.906	100
5	0.272	0.439	0.613	**0.442**	**35**	**34**	0.694	0.818	0.911	0.808	97
6	0.558	0.713	0.855	0.709	89	**35**	0.247	0.361	0.508	**0.372**	**51**
7	0.275	0.405	0.582	**0.421**	**42**	**36**	0.650	0.784	0.895	0.776	89
8	0.514	0.658	0.808	0.660	89	**37**	0.636	0.763	0.884	0.761	86
9	0.839	0.926	0.976	0.914	100	**38**	0.178	0.287	0.447	**0.304**	**37**
10	0.775	0.874	0.961	0.870	97	**39**	0.769	0.863	0.926	0.853	100
11	0.342	0.468	0.632	**0.481**	**59**	**40**	0.714	0.832	0.937	0.827	100
12	0.644	0.766	0.868	0.760	88	**41**	0.681	0.789	0.889	0.787	96
13	0.753	0.858	0.929	0.847	90	**42**	0.278	0.429	0.600	**0.436**	**61**
14	0.494	0.647	0.784	0.642	88	**43**	0.544	0.697	0.853	0.698	80
15	0.233	0.382	0.553	**0.389**	**60**	**44**	0.183	0.321	0.497	**0.334**	**46**
16	0.750	0.863	0.939	0.851	100	**45**	0.456	0.621	0.782	0.619	84
17	0.658	0.779	0.884	0.774	92	**46**	0.406	0.584	0.768	0.586	77
18	0.211	0.353	0.534	**0.366**	**36**	**47**	0.381	0.561	0.761	0.567	78
19	0.686	0.808	0.905	0.800	94	**48**	0.189	0.332	0.508	**0.343**	**35**
20	0.653	0.795	0.903	0.783	90	**49**	0.606	0.739	0.853	0.733	89
21	0.739	0.845	0.921	0.835	98	**50**	0.228	0.379	0.568	**0.392**	**28**
22	0.681	0.803	0.905	0.796	90	**51**	0.144	0.261	0.421	**0.275**	**25**
23	0.714	0.821	0.913	0.816	88	**52**	0.647	0.787	0.908	0.781	90
24	0.111	0.211	0.382	**0.234**	**31**	**53**	0.664	0.803	0.921	0.796	93
25	0.686	0.805	0.903	0.798	89	**54**	0.706	0.834	0.934	0.825	100
26	0.561	0.682	0.795	0.679	84	**55**	0.528	0.703	0.871	0.700	89
27	0.772	0.876	0.945	0.864	90	**56**	0.631	0.782	0.921	0.778	90
28	0.700	0.816	0.911	0.809	91	**57**	0.633	0.771	0.905	0.770	88
29	0.700	0.813	0.916	0.810	92						

For any item in FDM analysis to be accepted, any of these three conditions should be met: a) the threshold value ≤ 0.2, b) the expert consensus percentage $\geq 75\%$, and c) the defuzzification value ≥ 0.5. As can be depicted in Table 4, 14 items were rejected as they have defuzzification values smaller than 0.5 and the frequency of the expert agreement was below 75%. As a result, 43 measures under five factors were confirmed for the first round of the Delphi method. Except for the deletion of 14 items, a thorough review of the comments of the experts also led to a few modifications in other items such as minor additions, deletion, change of a word, an item rewording, or merging part of a deleted item in the existing item. With these changes in the remaining items and to have more refinement and confirmation of the existing results, the researchers decided to send the remaining factors (N=5) and measures (N=43) to the panel members again. It is noteworthy that an Excel file including the aggregate results of all the panel members' responses and comments to every individual item remaining from round 1 was also anonymously sent to the panelists to establish a new consensus regarding the existing items.

Table 5. The results of round 2 on the proposed model

Measures	Triangular Fuzzy Numbers			Defuzzi-fication Value	Threshold Value	Measures	Triangular Fuzzy Numbers			Defuzzi-fication Value	Threshold Value
	m1	m2	m3				m1	m2	m3		
1	0.792	0.929	0.995	0.905	0.02	**29**	0.736	0.845	0.942	0.841	0.03
2	0.644	0.774	0.897	0.772	0.03	**30**	0.642	0.808	0.939	0.846	0.15
3	0.722	0.855	0.961	0.846	0.02	**32**	0.775	0.887	0.971	0.878	0.08
4	0.742	0.855	0.945	0.847	0.02	**33**	0.861	0.945	0.995	0.934	0.03

6	0.628	0.784	0.905	0.772	0.06	**34**	0.711	0.839	0.934	0.828	0.02
8	0.536	0.682	0.829	0.682	0.02	**36**	0.747	0.876	0.974	0.893	0.12
9	0.878	0.955	0.995	0.943	0.03	**37**	0.744	0.868	0.966	0.878	0.12
10	0.811	0.908	0.982	0.900	0.03	**39**	0.833	0.929	0.984	0.915	0.06
12	0.650	0.776	0.879	0.768	0.01	**40**	0.811	0.908	0.982	0.900	0.07
13	0.758	0.868	0.939	0.855	0.01	**41**	0.686	0.800	0.900	0.795	0.01
14	0.506	0.658	0.797	0.654	0.01	**43**	0.642	0.795	0.929	0.840	0.14
16	0.825	0.913	0.968	0.902	0.05	**45**	0.489	0.655	0.816	0.653	0.03
17	0.767	0.882	0.971	0.887	0.10	**46**	0.742	0.874	0.968	0.861	**0.28**
19	0.778	0.889	0.966	0.878	0.08	**47**	0.689	0.829	0.947	0.822	**0.25**
20	0.744	0.876	0.963	0.861	0.08	**49**	0.878	0.955	0.995	0.943	**0.21**
21	0.761	0.868	0.942	0.857	0.02	**52**	0.683	0.821	0.929	0.811	0.03
22	0.753	0.868	0.953	0.858	0.06	**53**	0.711	0.850	0.950	0.837	0.04
23	0.819	0.913	0.982	0.905	0.09	**54**	0.706	0.834	0.934	0.825	0.00
25	0.775	0.895	0.968	0.879	0.08	**55**	0.539	0.716	0.879	0.711	0.01
26	0.578	0.703	0.818	0.700	0.02	**56**	0.653	0.808	0.937	0.799	0.02
27	0.808	0.908	0.971	0.896	0.03	**57**	0.683	0.813	0.932	0.809	0.04
28	0.731	0.855	0.947	0.844	0.04						

Table 6. Expert agreement consensus on the factors and measures – Round 2

Factors	**Measures**	**Expert**
Cognitive	1. I can focus on two or more different tasks simultaneously. (e.g., doing calculations while I am listening to a news broadcast).	95
Cognitive	2. I can switch my focus back and forth between tasks with different cognitive demands. (e.g., watching movies and helping my brother with homework).	92
Cognitive	3. I can focus on one specific task for a long time without being distracted.	97
Cognitive	4. When I work on a difficult project, I often make careless mistakes.	97
Cognitive	6. I have difficulty in planning and completing a detailed task that requires organization.	85
Cognitive	8. My thoughts (worriedness, fear of failure, daydreaming) distract my concentration.	96
Cognitive	9. I can return full attention to something after a short interruption.	91
Cognitive	10. I can concentrate even if I have little time to finish the task.	92
Cognitive	12. I can easily concentrate even if the subject matter is very complex or difficult.	98
Cognitive	13. When I have distracting thoughts, I step back without getting taken over by them.	100
Cognitive	14. I can focus on the present moment without my mind being preoccupied with the future or the past.	100
Cognitive	16. I tend to focus on one thing at a time rather than doing several things at once.	87
Cognitive	17. I increase my concentration by daily reflection on my actions.	78
Socio-emotional	19. I can concentrate since I often try to be relaxed.	80
Socio-emotional	20. If I am not sufficiently motivated or engaged in a task, I can manage to keep my attention.	79
Socio-emotional	21. I can control my attention by changing the way I think about the situation I'm in.	96
Socio-emotional	22. In stressful situations, I make myself think about it in a way that helps me stay calm.	85
Socio-emotional	23. I can effectively identify my emotions and thoughts as well as their impact on my actions.	80
Technological	25. I need to limit my screen time (e.g., use of mobile phone, tablet, internet, & social media) to enhance my attention.	81
Technological	26. My screen time (e.g., use of mobile phone, tablet, internet, & social media) does not distract my thoughts.	96
Technological	27. When using an application, having a reminder to close all other apps can help focus my attention on the task at hand.	92
Technological	28. I can keep a balanced attention between my daily activities/tasks and screen time.	90
Technological	29. I can manage any distracting challenges (e.g., receiving calls, messages, notifications, ...) caused by my screen use.	92

Technological	30. I find it difficult to concentrate in an online class.	77
Environmental	32. I am distracted by events or noise around me.	80
Environmental	33. Changes in my physical space (e.g., changing one's school) reduce my concentration.	90
Environmental	34. It distracts my attention if an examiner comes and stands near me while I am responding to the exam items.	97
Environmental	36. I cannot concentrate when I am cold/hot, hungry/thirsty, or sleepless.	78
Environmental	37. The physical space of the environment must be quiet for me to concentrate or study effectively.	79
Environmental	39. Studying in a group reduces my concentration.	88
Environmental	40. A safe learning environment in which I can trust the process and the learning community (e.g., teachers and classmates) helps me be more focused.	84
Environmental	41. Inappropriate classroom equipment affects the level of my attention.	100
Environmental	43. Family problems can affect my concentration.	76
Environmental	45. Economic problems can be a source of my inattentiveness in the classroom.	90
Socio-emotional	**46. Academic achievement is an important factor in increasing my attention in the classroom.**	**60**
Socio-emotional	**47. Overcoming distracting challenges of everyday academic life (e.g. test anxiety, time management, class workload, …) can enhance my focus.**	**66**
Environmental	**49. Being affected by friends or competing against them can be a source of my inattention.**	**70**
Pedagogical	52. I try to be more focused by following my teacher's effective feedback.	90
Pedagogical	53. Lack of variety in classroom activities cannot distract me.	88
Pedagogical	54. I need deadlines to be able to be more focused on my tasks.	100
Pedagogical	55. My concentration is not reduced when the teaching methods do not fit my learning style.	100
Pedagogical	56. I keep focused during my study even if the break time is insufficient.	96
Cognitive	57. Daily meditation before the class/studying time enhances my concentration.	88

Table 6 shows the results of FDM analysis based on the second round of responses. As can be seen in the table, all measures that have threshold values below 0.2 received the approval of the expert panel. However, measures 46, 47, and 49 have threshold values above 0.2 which have not met the condition and are rejected. As displayed in Table 4, the expert agreement on the measures is also indicative of the low consensus ($\leq 75\%$) on the above three measures. Hence, they can be deleted from the scale.

Table 7. Dimensions, factors, measures and expert agreement

Dimensions	Factors	No. of Measures	Expert Agreement
Internal	Cognitive	14	92.57
	Socio-emotional	5	84
External	Technological	6	88
	Environmental	10	86.2
	Pedagogical	5	94.8

To investigate the position of each dimension, factor, and measure in the model among the expert panel members, the results of aggregating the expert agreements on the dimensions of the attentional literacy scale in Table 7 indicate that experts have almost equal consensus on internal (88.28%) and external (89.66%) dimensions of the concept. Furthermore, pedagogical (94.8%) and cognitive (92.57%) factors are the most agreed factors while socio-emotional factor (84%) is the least agreed factor. Table 7 also shows the position of each measure in the model.

5. Discussion

This study was intended to develop an attentional literacy model for language learners. Using the Fuzzy Delphi technique, the researchers aimed to seek a level of consensus among the experts in the related fields regarding the dimensions, factors, and measures in this model. The results of two rounds of analysis led to the transformation of the initial proposed model. The revised attentional literacy model includes two dimensions, five factors, and 40 measures. There are two dimensions: internal and external. The internal dimension of the attentional literacy framework refers to resources that initiate from within an individual that can represent the level of attentional literacy in a person. It can help or hinder attention management. This dimension includes two factors: cognitive and socio-emotional. The cognitive factor entails executive functions which include top-down mental processes necessary to concentrate and pay attention. These functions are response inhibition, working memory, and flexible thinking. It also includes sustained attention which refers to the ability to maintain attentional focus over extended periods. The socio-emotional factors are those that influence a person to understand, experience, express, and manage emotions (emotional) and to develop meaningful relationships with others (social).

The external dimension refers to the resources which reside outside of a person either in the physical or social contexts of a person. They can reinforce or interfere with the attentional capacity of the learners. The external dimension is comprised of three factors technological, pedagogical, and environmental. Technological factor in our attentional literacy framework concerns the use of and effect of digital devices, such as individual screen time or attending virtual classes. The pedagogical factor entails those related to the instructional context of the classroom like teacher-related issues (teaching methods, types of feedback, …), instructional activities/tasks, and materials that can affect the level of attention. Environmental factors are those originating from the immediate surroundings such as space, location, shape, color, lighting, and sounds that can affect the attention.

Figure 3 shows the schematic representation of the final version of the attentional literacy model.

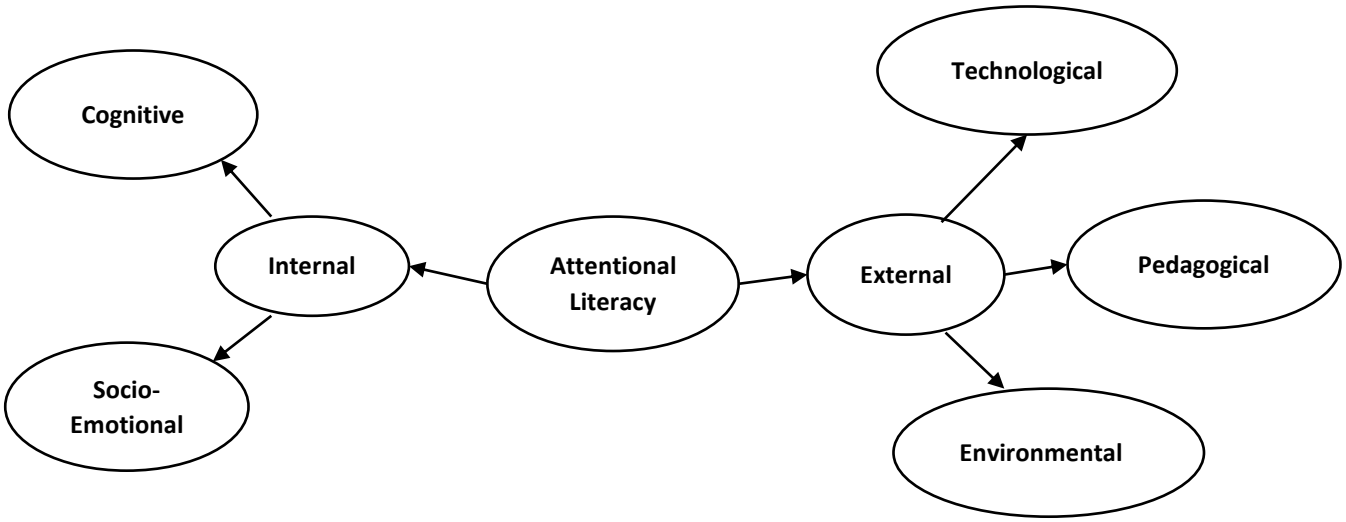

Figure 3. The final version of the attentional literacy conceptual model

The interrelatedness of the components in the model is a significant finding that underscores the complexity of attentional literacy. That is, these components do not stand alone within the external and internal dimensions. On the contrary, they are in a dynamic interplay and mutually influencing learners' attentional capacities. In other words, the students' cognitive ability to concentrate is the result of their emotional state as well as their teacher's pedagogical methods adopted in class. Similarly, the external technological environment like (not) having notifications and text messages on digital gadgets while studying can promote or impede cognitive and socio-emotional factors. This network of interconnection among the components of attentional literacy suggests that students, educators, and curriculum designers should employ a holistic approach to adopt and adapt strategies in language classes which encompasses both internal and external factors. On the whole, the findings of this study expand the theoretical foundations of the concept of attentional literacy by proposing a model and the underlying dimensions, factors, and measures. Literature lacks a multifaceted model for attentional literacy although a few studies have been conducted on some of the indicators of the present study. The findings of this study revealed that the pedagogical factor, among others, was the most agreed-upon factor in the attentional literacy model. This corroborates the findings of other studies which proposed that the practices of the teachers in classrooms have an impact on the students' attention (Dao et al., 2020; Hlas, et al., 2019; Wang, 2015) and reported strategies for teachers to develop learners' attentional literacy (Palalas et al., 2024). The role of other factors of the newly proposed model was also reported separately in the literature as pivotal in the learners' attentional literacy. Therefore, our findings are in line with studies by Schepers (2007) for the environmental influences, Yip (2023a) and Yip et al., (2023b) for the cognitive and emotional dimensions of attention, and Pegrum and Palalas (2021) and Palalas et al. (2024) for the learners' digital disarrays and hyperconnectivity which are effective in maintaining and directing the attention.

6. Conclusion

This study adopted the FDM to explore the dimensions, factors, and measures of the attentional literacy model. After examining the related literature, designing the factors and measures of an initial model, selecting the expert panel members, and administering the measures in two rounds, the data from the responses and the comments were analyzed. The results offer several contributions to the field. First, the use of FDM is not so prevalent in language education. Yet, when it is used in studies, like the current one, it proved to be efficient in orchestrating the conflicting views of the experts to achieve a desired consensus. The method also offers an effective methodological approach to develop and validate a theoretical model. The results of this study also contribute to the present literature by introducing a model to measure attentional literacy among TEFL students. The model has two internal and external dimensions with five factors and 40 measures which operationalize the notion of attentional literacy from cognitive, socio-emotional, environmental, pedagogical, and technological perspectives. Viewed through the lens of the sociocultural paradigm, the concept is multi-faceted thus selecting experts from different subdisciplines of Applied Linguistics and general education added to the value of the resulting multidisciplinary model. AL was seen as a cognitive and pedagogical issue with almost half of the measures falling into those categories which achieved a high level of agreement among the experts.

The other half spread among the technological, environmental, and socio-emotional issues for language learners which achieved even higher level of agreement. Interestingly enough, the pedagogical factor received the highest level of consensus among the experts indicating that effective teaching practices are paramount in promoting attentional literacy among language learners.

The insights from the study have certain implications. The principal theoretical implication of the study is that the concept of attentional literacy is multidimensional and many factors are involved to support the language learners' attentional capacity. This requires employing a holistic approach to address all factors contributing to the learners' attentional literacy. In terms of practical implication, all the stakeholders like educators, language school administrators, and policy-makers are suggested to foster environments in which the cognitive strategies are integrated in the classroom activities, socio-emotional skills are empowered, technological literacy is mindfully addressed, diverse pedagogical methods to cater for different learning styles are enhanced, and supportive learning environments are to be created.

Despite all the effort exerted to the meticulous implementation of the research, there are a few limitations that can impact the generalizability and accuracy of the results. As mentioned in the significance of the study, there was limited literature regarding the AL concept within the English language education which compelled the researchers to delve into the related literature within the field of general education. This can reduce the specificity of the concept. Also including more panel members could enhance the accuracy and generalizability of the model and more varied views could be proposed. Furthermore, the subjective nature of expert opinions might threaten the validity of the research findings. In this regard, we recommend testing the model with language learners to further improve its validity. Moreover, it is recommended that the current model be further reconceptualized in other educational settings to enhance its robustness and validity.

7. References

Al-Rikabi, Y. K., & Montazer, G. A. (2024). Designing an e-learning readiness assessment model for Iraqi universities employing Fuzzy Delphi Method. *Education and Information Technologies*, *29*(2), 2217-2257. https://doi.org/10.1007/s10639-023-11889-0

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Washington, DC.

Arnold, L. E., Hodgkins, P., Kahle, J., Madhoo, M., & Kewley, G. (2020). Long-term outcomes of ADHD: Academic achievement and performance. *Journal of Attention Disorders*, *24*(1), 73-85. https://doi.org/10.1177/1087054714566076

Berger, J., Moe, W. W., & Schweidel, D. A. (2023). What holds attention? Linguistic drivers of engagement. *Journal of Marketing*, *87*(5), 793-809. https://doi.org/10.1177/00222429231152880

Bonnardel, V., Biddington, T., May, B., Jones, R., & Roffey, S. (2018). Toward the implementation of contemplative practices in higher education. *Journal of Perspectives in Applied Academic Practice*, *6*(3), 3-13. https://doi.org/10.14297/jpaap.v6i3.362

Bulusan, F. (2024). An emerging instructional materials design model for English as a second language (ESL) and English as a foreign language (EFL) senior high school contexts. *Innovation in Language Learning and Teaching*, 1–14. https://doi.org/10.1080/17501229.2024.2431890

Burgoyne, A. P., & Engle, R. W. (2020). Attention control: A cornerstone of higher-order cognition. *Current Directions in Psychological Science*, *29*(6), 624-630. https://doi.org/10.1177/0963721420969371

Chang, P. T., Huang, L. C., & Lin, H. J. (2000). The fuzzy Delphi method via fuzzy statistics and membership function fitting and an application to the human resources. *Fuzzy Sets and Systems*, *112*(3), 511-520. https://doi.org/10.1016/S0165-0114(98)00067-0

Chapman, M. (1987). Piaget, attentional capacity, and the functional implications of formal structure. *Advances in Child Development and Behavior*, *20*, 289-334. https://doi.org/10.1016/S0065-2407(08)60405-X

Dao, P., Chi Nguyen, M. X. N., & Chi, D. N. (2020). Reflective learning practice for promoting adolescent EFL learners' attention to form. *Innovation in Language Learning and Teaching*, *15*(3), 247–262. https://doi.org/10.1080/17501229.2020.1766467

Dong, Y., & Li, P. (2020). Attentional control in interpreting: A model of language control and processing control. *Bilingualism: Language and Cognition*, *23*(4), 716-728. https://doi.org/10.1017/S1366728919000786

Dörnyei, Z. (2020). *Innovations and challenges in language learning motivation*. Routledge. https://doi.org/10.4324/9780429485893

Dörnyei, Z., & Ushioda, E. (2021). *Teaching and researching motivation*. Routledge.

Ellis, R. (2008). *The study of second language acquisition*. OUP.

Español-Martín, G., Pagerols, M., Prat, R., Rivas, C., Ramos-Quiroga, J. A., Casas, M., & Bosch, R. (2023). The impact of attention-deficit/hyperactivity disorder and specific learning disorders on academic performance in Spanish children from a low-middle-and a high-income population. *Frontiers in Psychiatry*, *14*, 1136994. https://doi.org/10.3389/fpsyt.2023.1136994

Eysenck, M. W. (2001). *Principles of cognitive psychology*. Psychology Press.

Eysenck, M. W., & Derakshan, N. (2011). New perspectives in attentional control theory. *Personality and Individual Differences*, *50*(7), 955-960. https://doi.org/10.1016/j.paid.2010.08.019

Finneran, D. A., Francis, A. L., & Leonard, L. B. (2009). Sustained attention in children with specific language impairment (SLI). *Journal of Speech, Language, and Hearing Research*, *52*(4), 915-929. https://doi.org/10.1044/1092-4388(2009/07-0053)

Ghani, M. T. A., Daud, W. A. A. W., Ramli, S., Hamat, M. Z. A., & Abdelhamid, I. Y. (2024). The Usability of Arabic Language Classroom Management Mobile Application Through Fuzzy Delphi Analysis. In *Artificial Intelligence (AI) and Customer Social Responsibility (CSR)* (pp. 989-1002). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-50939-1_80

Gray, S. A., Dueck, K., Rogers, M., & Tannock, R. (2017). Qualitative review synthesis: The relationship between inattention and academic achievement. *Educational Research*, *59*(1), 17-35. https://doi.org/10.1080/00131881.2016.1274235

Griffiths, C. (2018). *The strategy factor in successful language learning: The tornado effect.* Multilingual Matters. https://doi.org/10.21832/9781783099757

Henry, A., & MacIntyre, P. D. (2024). *Willingness to communicate, multilingualism and interactions in community contexts* (Vol. 22). Multilingual Matters.

Hlas, A. C., Neyers, K., & Molitor, S. (2019). Measuring student attention in the second language classroom. *Language Teaching Research*, *23*(1), 107-125. https://doi.org/10.1177/1362168817713766

Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching*, *43*(2), 154-167.

Issa, B. I., & Morgan-Short, K. (2019). Effects of external and internal attentional manipulations on second language grammar development: An eye-tracking study. *Studies in Second Language Acquisition*, *41*(2), 389-417.

James, W. (1890). *The principles of psychology*. Henry Holt.

Komorowska, H. (2021). The role of attention in teacher education: A Factor in the quality of European schooling. *Theory and Practice of Second Language Acquisition*, *1*(7), 33-49.

Krashen, S. (1981). Second language acquisition. *Second Language Learning*, *3*(7), 19-39.

Liu, Y.-T., Nassaji, H., & Tseng, W.-T. (2024). Effects of internal and external attentional manipulations and working memory on second language vocabulary learning. *Language Teaching Research*, *28*(4), 1701-1741. https://doi.org/10.1177/13621688211030130

Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie & T. K. Bhatia (Eds.), *Handbook of Second Language Acquisition (Vol. 2)* (pp. 413- 468). Academic Press.

Luo, M., DeWitt, D., & Alias, N. (2024). Experts' consensus on the design of a technical communication module for vocational college students in the manufacturing sector. *Language Teaching Research.* https://doi.org/10.1177/13621688231223819

MacIntyre, P. D., & Wang, L. (2021). Willingness to communicate in the L2 about meaningful photos: Application of the pyramid model of WTC. *Language Teaching Research*, *25*(6), 878-898. https://doi.org/10.1177/13621688211004645

Manakandan, S. K., Rosnah, I., Mohd, R. J., & Priya, R. (2017). Pesticide applicators questionnaire content validation: A fuzzy Delphi method. *Medical Journal of Malaysia*, *72*(4), 228-235.

Marlina, E., Hidayanto, A. N., & Purwandari, B. (2022). Towards a model of research data management readiness in Indonesian context: An investigation of factors and indicators through the fuzzy Delphi method. *Library & Information Science Research, 44*(1), 101141. https://doi.org/10.1016/j.lisr. 2022.101141

Martínez-Vicente, M., Martínez-Valderrey, V., Suárez-Riveiro, J. M., & Valiente-Barroso, C. (2023). Relationship between learning English as a foreign language and the executive attention profile in Spanish schoolchildren. *Psicología Educativa. Revista de los Psicólogos de la Educación*, *29*(2), 159-166. http://dx.doi.org/10.5093/psed2023a10

Melo, C., Madariaga, L., Nussbaum, M., Heller, R., Bennett, S., Tsai, C.-C., & van Braak, J. (2020). Editorial: Educational technology and addictions. *Computers & Education, 145*. https://doi.org/10.1016/j.compedu.2019.103730

Melo, J. E. A., Ariza, H. H. L., & Marin, E. V. L. (2024). Physical activity in students and its association with attention: A systematic review. *Retos: nuevas tendencias en educación física, deporte y recreación*, (56), 834-845.

Miyadera, C. (2024). The impact of inattention, hyperactivity/impulsivity symptoms, academic achievement, and gender on the self-esteem: A study of children in Japanese elementary schools. *Cogent Psychology*, *11*(1), 2354961. https://doi.org/10.1080/23311908.2024.2354961

Mohebbi, A. (2023). Optimizing Language Learning for Students with ADHD: Strategies for Cultivating Attentional Mechanisms. In X. Jiang (Ed.), *Psycholinguistics: New Advances and Real-World Applications* (pp. 134- 145). IntechOpen. https://doi.org/10.5772/intechopen.1003025

Nurmuhammad, A. (2023). Features of cognitive processes such as perception, memory, thinking in students of junior school age. *Education News: Research in the 21st Century, 1*(10), 785–791.

Ott, E. (1994). *Das konzentration spogramm. Konzentrationsschwaeche-überwinden-denkvermöqen steigen* [The concentration program. Overcoming poor concentration and improving thinking skills]. Rowohlt Taschenbuch Verlag.

Oxford, R. L. (2017). *Teaching and researching language learning strategies: Self-regulation in context* (2nd ed.). Routledge. https://doi.org/10.4324/9781315719146

Palalas, A., Pegrum, M., & Dell, D. (2024). Regaining focus: Promoting attentional literacy in digital higher education. *Australasian Journal of Educational Technology*. https://doi.org/10.14742/ajet.9627

Pegrum, M., & Palalas, A. (2021). Attentional literacy as a new literacy: Helping students deal with digital disarray. *Canadian Journal of Learning and Technology*, *47*(2), 1-18. https://doi.org/10.21432/cjlt28037

Robinson, P. (2003). Attention and memory during SLA. In C. J. Doughty, & M. H. Long (Eds.), *The Handbook of Second Language Acquisition* (pp. 631 – 78). Malden, MA: Blackwell. https://doi.org/10.1002/9780470756492

Rogoff, B. (1990). *Apprenticeship in thinking: Cognitive development in social context*. Oxford University Press.

Schepers, J. M. (2007). The construction and evaluation of an attention questionnaire. *SA Journal of Industrial Psychology*, *33*(2), 16-24.

Schmidt, R. (1994). Deconstructing consciousness in search of useful definitions for applied linguistics. *Consciousness in Second Language Learning*, *11*, 237-326.

Sensuse, D. I., Purwandari, B., & Rahayu, P. (2018). Defining e-Portofolio Factor for Competency Certification using Fuzzy Delphi Method. *Turkish Online Journal of Educational Technology-TOJET, 17*(2), 25-33. https://eric.ed.gov/?id=EJ1176167

Sterling, S., Plonsky, L., Larsson, T., Kytö, M., & Yaw, K. (2023). Introducing and illustrating the Delphi method for applied linguistics research. *Research Methods in Applied Linguistics*, *2*(1), 100040. https://doi.org/10.1016/j.rmal.2022.100040

Stevens, E. A., Stewart, A., Vaughn, S., Lee, Y. R., Scammacca, N., & Swanson, E. (2024). The effects of a tier 2 reading comprehension intervention aligned to tier 1 instruction for fourth graders with inattention and reading difficulties. *Journal of School Psychology*, *105*, https://doi.org/10.1016/j.jsp.2024.101320

Stewart, A. A., & Swanson, E. (2024). Supporting reading comprehension for students with inattention. *Teaching Exceptional Children*, *56*(5), 386-395. https://doi.org/10.1177/00400599221087434

Taghavi-Nejad, F. S., Fallah, N., & Lotfi Gaskaree, B. (2024). Mindfulness and procrastination among university EFL learners: The role of attention control and self-regulated learning. *Psychological Reports*, *0*(0). https://doi.org/10.1177/00332941241287423

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *41*(2), 363-387. https://doi.org/10.1017/S0272263118000311

Tomlin, R. S., & Villa, V. (1994). Attention in cognitive science and second language acquisition. *Studies in Second Language Acquisition*, *16*(2), 183-203. https://doi.org/10.1017/S0272263100012870

VanPatten, B. (1996). *Input processing and grammar instruction: Theory and research*. Ablex.

Vygotsky, L. S. (1994). The problem of the environment. In R. van der Veer & J. Valsiner (Eds.), *The Vygotsky Reader* (pp. 338–354). Blackwell.

Wang, W. (2015). Factors affecting learners' attention to teacher talk in nine ESL classrooms. *TESL-EJ, 19*, 1–20.

Yasin, A. A., Masri, R., & Adnan, M. (2022). Application of fuzzy delphi method in designing a basic teaching model based on moral values in stem mathematical. *Journal of Positive School Psychology*, *6*(3), 4898-4920.

Yip, J. M. Y. W. (2023). *Cognitive and affective dimensions of inattention*. Doctoral dissertation, University of British Columbia.

Yip, J. M., Jodoin, N. M., & Handy, T. C. (2023). Dimensions of inattention: Cognitive, behavioral, and affective consequences. *Frontiers in Psychology*, *14*, 1075953. https://doi.org/10.3389/fpsyg.2023.1075953

Zadeh, L. A. (1965). Fuzzy sets. *Information and Control, 8*(3), 338-353. https://doi.org/10.1016/S0019-9958(65)90241-X

Ziaei, S., & Dabaghi, A. (2023). Effects of corrective feedback on Iranian EFL learner's uptake and attention. *Applied Linguistics Inquiry, 1*(1), 129-141. https://doi.org/10.22077/ali.2023.6268.1013

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 73-86

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.9031.1109

Impact of Employing Photovoice Integrated with Creative Problem Solving Model on Iranian Intermediate EFL Learners' Speaking Skill and Attitudes

Somayeh Pournia[1]
Azizeh Chalak[2*]
Hossein Heidari Tabrizi[3]

[1]Ph.D. Candidate in Applied Linguistics, Department of English, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran
[*2,3]Professor of Applied Linguistics, Department of English, Isfahan (Khorasgan) Branch, Islamic Azad University, Isfahan, Iran

ARTICLE HISTORY

Received: 20 December 2024
Revised: 12 January 2025
Accepted: 17 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: azizeh.chalak@iau.ac.ir

ABSTRACT

Speaking is a multifaceted skill that involves many different areas, such as discourse, grammar, and sociolinguistics. As a primary means of communication, speaking is essential in any instructional program for language learning. Besides, Photovoice is a participatory learning method that represents the viewpoints, issues, and experiences of individuals or communities through photography. Following a mixed-methods design, the present study tried to investigate employing Photovoice on Iranian intermediate language learners' skill in speaking English. It also looked at how the participants felt about using Photovoice to learn languages. Using a convenient sampling procedure, a total number of 60 participants were chosen. Once their homogeneity was confirmed based on the Preliminary English Test, they were randomly divided into two equal groups of thirty participants, one experimental group and one control group. While the control group received typical language teaching using conventional methods, the experimental group followed the PHOTO model, which emphasized participatory photography, as well as the Creative Problem Solving (CPS) model for generating innovative solutions. Triangulation was also accomplished through conducting pre/post tests, a questionnaire, and an interview. Then, the data were summarized, and descriptive and inferential statistics such as one-way ANCOVA and paired samples t-tests were used. The results showed that Photovoice integrated with CPS was an effective method for enhancing students' speaking skill and emotional engagement in language learning. The findings can be helpful through new insights into language pedagogy, suggestions for researchers and language teachers in the study of speaking skill, and visual methods such as Photovoice.

KEYWORDS: Creative Problem Solving (CPS) model; Participatory learning; PHOTO model; Photovoice; Speaking skill

1. Introduction

Speaking skill plays a significant role in interpersonal communication, and improving learners' speaking abilities appears to be necessary. The majority of language learners evaluate their proficiency in a language primarily based on their speaking skill. Students should be able to communicate what they mean and comprehend what others mean to strengthen their language skills. Besides, many teachers believe that the purpose of teaching speaking is to enable students to effectively communicate by allowing them to express their opinions, feelings, and thoughts.

Cite this article: Pournia, S., Chalak, A. and Heidari Tabrizi, H. (2025). Impact of employing photovoice integrated with creative problem solving model on Iranian intermediate EFL learners' speaking skill and attitudes. *Applied Linguistics Inquiry*, *3*(1), 73-86. doi: 10.22077/ali.2025.9031.1109

Photovoice is built upon the participatory methods of documentary photography, which provide context for individuals to express their perspectives (Nisa, 2021). Through the use of images, Photovoice participants can readily communicate their experiences and thoughts. This method has been recognized for promoting interactive learning and developing students' innovative thinking (Ferdiansyah et al., 2020), as well as boosting students' oral proficiency in a secondary English-specialized class (Bhatti, 2021). Likewise, teachers' questions can be accompanied by the pictures. The interactive aspect of the questions that the teachers provide helps to clarify the connection between the methods and the students' learning growth (Sa'adah et al., 2023).

A variety of teaching methods has been used to improve students' speaking skill, yet classroom observations and research have demonstrated that plenty of students have difficulties with speaking skill (Ashimovna, 2022). Based on Ur's (2021) classification, problems include limited involvement, inhibition, the use of the mother tongue in the classroom, and the absence of a theme to be spoken. Speaking in front of the class can be a source of anxiety for language learners. In addition to having trouble choosing a topic to discuss, some people struggle to articulate themselves in front of the teacher. Finding the right moment to contribute to debates can be difficult for certain other students, and some of them tend to use their mother tongue. These issues hinder language learners' ability to speak accurately and fluently.

In light of the above discussion, the difficulties that foreign language learners experience while speaking have led researchers to carry out the study on using the Photovoice method to teach speaking in the hopes of resolving the problems by incorporating learning media into learners' reflective learning. Despite the growing emphasis on communicative language teaching, EFL learners in Iran still struggle with speaking. For this reason, additional methods that actively engage students and improve their communication skills are needed. As a participatory learning method, Photovoice has obvious advantages over conventional teaching methods, since it allows students a reflective and pictorial means of communicating ideas, leading to greater engagement and less fear of speaking English in public. Through the integration of visual prompts, Photovoice creates a more dynamic and engaging learning environment in comparison to traditional methods, which largely rely on artificial speech practice.

Additionally, this study is significant because of its theoretical, practical, and pedagogical insights into language learning. Theoretically, it introduces Photovoice and creative problem solving (CPS) into the language learning methods, with implications on how visual-based participatory methodology constructs speaking development. Practically, Photovoice can be applied to language education to enrich language learning by ensuring participant involvement in meaningful communicative activities. It can provide EFL teachers with an innovative methodology to decrease the monotonous and boring atmosphere of speaking classes. Pedagogically, curriculum designers can strike a balance between procedural and theoretical knowledge. They may include interesting content to the curriculum and modify English speaking process to make it more dynamic, learner-centered, and process-oriented.

Moreover, utilizing Photovoice in the context of intermediate learners in Iran has not been extensively explored (Karimi et al. 2018, 2019; Pournia et al., 2025). It highlights the need for action research employing Photovoice to evaluate its effectiveness in addressing students' speaking issues. Thus, in the context of Iranian EFL, the research objectives were to investigate the Photovoice method integrated with CPS on the students' speaking skill and attitudes considering their experiences.

2. Literature review

The following sections synthesize a variety of sources to discuss theoretical and empirical research on Photovoice and English language teaching and learning.

2.1. Theoretical background of the study

Speaking is generally a skill for people to communicate with one another (Bhatti, 2021). It has been evident that students who can express their concepts and ideas via language in an impressive manner are higher achievers in academics (Ashimovna, 2022). Several methods have been proposed to eliminate learners' issues in speaking classes by managing both the quantity and quality of interactions in the speaking classes (Purti, 2018; Sa'adah et al., 2023; Sarani et al., 2020). Moreover, Photovoice is a participatory learning method that can be utilized to bring about individual and collective transformation (Ordem, 2023). Teachers' questions can be used to guide students to reflect on their learning process and regulate how involved they are in speaking activities (Banitalebi & Ghiasvand, 2023).

A review of the literature revealed that four primary Photovoice protocols or models had been used as a framework for speaking a language. Drawing upon Wang and Burris's (1994) SHOWeD model, students elaborated on the story's context, its characters, time and place, and the initial situation. Based on Graziano's (2004) Photovoice model, each photograph had a set of guiding questions. Participants explained the picture, what was happening in it, and why they took it, helping them explore their English language learning experiences and how the photos could create opportunities for improvement. In Koltz et al.'s (2010) PHOTO model, learners discussed their photos, which, combined with prompt reminders, led to powerful storytelling. They described the photo’s context, their motivations for taking it, and their emotions while speaking. Finally, using Mitchell et al.'s (2018) 3Ws Photovoice model, students started by describing what they noticed in the picture, then explained how and why the phenomenon occurred, and why they took the picture. Answering the 3Ws model questions encouraged them to generate possible responses.

2.2. Theoretical framework of the study

In this study, the researchers followed Mitchell et al.'s (2018) 3Ws Photovoice model, a participatory learning method, as well as Treffinger et al.'s (2023) Creative Problem Solving (CPS) model, a model for generating and executing creative solutions systematically. Their inclusion is theoretically grounded in Sociocultural Theory (Vygotsky, 1978), Experiential Learning Theory (Kolb, 1984), and Communicative Language Teaching (CLT) that offer justification for interactive, student-centered teaching methods. The steps included the introduction of Photovoice in which participants were oriented to the process; understanding the challenge by determining the most significant issues to resolve; idea generation via brainstorming; preparation for action by organizing and planning for resources; taking action by implementing projects; reflecting and evaluating by determining outcomes and obtaining feedback; and finally, presenting a gallery walk to communicate the findings and engage with the visitors.

2.3. Empirical studies on photovoice in language learning and teaching

Photovoice can be used successfully in educational reform in addition to being a helpful method for researching social activities (Ciolan & Manasia, 2024; Sa'adah et al., 2023; Wass et al., 2020). Wass et al. (2020) combined more conventional data-collecting methods, such as open-ended focus group questions and the critical incident technique, with Photovoice, a visual data-collection method. They concluded that Photovoice provided a valuable methodological supplement to High Education (HE) research. However, they also pointed out that researchers should be aware of its limits. Ciolan and Manasia (2024) introduced self-written narratives and photos as data sources. The study demonstrated that offering the opportunity for active involvement in the learning process makes innovatory education the leading movement for teachers, institutions, and learners.

Furthermore, the Photovoice method sought to document and reflect students' strengths and concerns regarding the ELT course, promote critical thinking, and influence ELT policy through both large- and small-group discussions of photos (Afifah et al., 2023). The process involved four steps, namely planning, pre-focus meetings, photo capturing, and reflecting (Boamah et al., 2022). Photovoice has been applied in a multitude of studies to engage learners in critical thinking, ignite their passion for creativity, and develop research skills (Ferdiansyah et al., 2020; Mudra et al., 2023; Nisa, 2021). Ferdiansyah et al. (2020) identified three critical dimensions among students: peer support, self-motivation, and assimilation, which led to sociocultural adaptation skill development. A qualitative study by Mudra et al. (2023) revealed that Photovoice helped EFL students express their intense emotions about producing research papers and the challenges of publication. Thus, Photovoice was a helpful method for students and teachers to understand and use self-reflection in writing assignments (Nisa, 2021).

It has been investigated in some studies to find out the qualities of effective EFL English teachers through the Photovoice meth (Mudra et al., 2023; Ordem, 2023; Sa'adah et al., 2023). They have stressed primary characteristics such as communication, cultural sensitivity, agility, dynamism, developing a joyful learning environment, and applying various teaching methods. The studies have revealed the importance of those parameters that successful EFL teachers have to possess (Sa'adah et al., 2023). Using the Photovoice method, the researchers examined the qualities of teachers from a different perspective, providing a more comprehensive understanding of what facilitates successful language teaching. Nonetheless, the effect of this method may differ due to diverse contexts and participants (Ordem, 2023).

Using Photovoice, and questioning strategies to enhance instructional materials in English language classes have been the subject of some studies in the Iranian EFL contexts (Banitalebi & Ghiasvand, 2023; Karimi et al., 2018, 2019). They have mainly concentrated on questioning strategies, teacher questions, and students' responses. Karimi et al. (2019) explored the effect of Photovoice on language learners' intercultural sensitivity and reflective thinking across gender. They found that Photovoice promoted learners' intercultural sensitivity and reflective thinking. In addition, the findings concluded that teachers should not limit their questioning to knowledge-based topics. They ought to pose questions that demand clarification and generate answers that are more in-depth and syntactically intricate (Banitalebi & Ghiasvand, 2023).

For EFL teachers in Iran, motivating students to collaborate in group activities that foster deeper engagement and comprehension has consistently posed a key challenge. Based on the findings of the relevant studies, cooperation has a notably beneficial effect on enhancing EFL learners' skills (Ghoushchi et al., 2021; Karimi et al., 2018; Sarani et al., 2020). Karimi et al. (2018) asserted that the utilization of Photovoice assignments in EFL classes could allow learners to investigate topics of personal interest and share creations with their classmates. The teachers' main focus should be on helping students become independent learners and fostering autonomy by teaching effective methods for learning languages (Sarani et al., 2020). When teachers have a well-designed teaching plan tailored to their students, academic success can be achieved, and learning English can become an enjoyable activity (Ghoushchi et al., 2021).

Regarding the aforementioned studies, while numerous studies have examined the effectiveness of photovoice in other academic domains, there aren't many studies that examine the effect of the Photovoice method integrated with the CPS model on the speaking abilities of Iranian intermediate EFL learners from an academic standpoint. Speaking is the skill to effectively use a language to communicate concepts, feelings, and information to and from other communicators. Given the advantages indicated above, it could be stated that Photovoice should be tried in classrooms. Students' speaking skill was expected to improve with this method, which could be used more regularly in the classroom. Furthermore, the researchers chose a setting where Photovoice was never used as a participatory learning method in language classes. Consequently, the following research questions were the focus of this study:

1. Does employing Photovoice as a participatory learning method integrated with CPS have any significant effect on Iranian intermediate EFL learners' speaking skill?

2. What are the Iranian intermediate EFL learners' attitudes toward employing Photovoice as a participatory learning method in their English language classes?

3. Methodology

The next section includes the research design and setting of the study, variables, participants, material and instruments, data collecting and data analysis procedures, as well as the explanations that address concerns about reliability and validity.

3.1. Design and context of the study

Using a convergent parallel mixed-methods design, the study collected both quantitative and qualitative data simultaneously. By collecting and analyzing both types of data separately but simultaneously, this design provided a comprehensive understanding of the impact of Photovoice on EFL learners' productive skills. The investigation followed a quasi-experimental design which comprised pretest and posttest equal groups. An experimental design made it easier for the researchers to evaluate how well Photovoice affected the speaking skill of language learners. Considering that Photovoice was a community-based language learning method, the photo-based studies relied heavily on community members, participants, and researchers (Karimi et al., 2019).

3.2. Participants

Using convenience sampling, out of 150 language learners enrolled in a private English language school in Ardestan, Iran, 60 intermediate EFL learners were selected. Because of the institutions' policies about interviewing, participant selection, and questionnaire distribution, individuals were chosen by convenient sampling. It was also chosen due to its practicality and ease of access; however, this method may limit the generalizability of the findings. Then, they were split up into two equal groups. Thirty individuals were allocated to the experimental group and another thirty to the control group. The participants have taken at least 8 years of English language classes. The scores of the Preliminary English Test (PET) were used to confirm the participants' homogeneity in English language proficiency, indicating that their proficiency was at an intermediate level. Table 1 briefly indicates the demographics of the participants.

Table 1. Demographic background of the participants

No. of Learners	60
Age	17-20 years
Gender	Female
Native Language	Persian
Target Language	English
Population	Iranian Intermediate EFL learners
Institute	Private English Language Institute, Ardestan
Sampling Technique	Convenience

In the present study, the researchers ensured the participants' respect, privacy rights, and protection from physical and psychological harm. All participants received information about the study and provided their consent for the data to be recorded and analyzed in compliance with ethical standards. They needed to provide written consent to participate in the research. More importantly, a digital copy of the images remained on the researchers' password-protected computer until data analysis was completed, and then it would be destroyed.

3.3. Material

Intermediate Top Notch 1B, authored by Saslow and Ascher (2015), served as the primary coursebook for both groups during the semester. Five units were designed to prepare language learners to communicate effectively in English, emphasizing all four skills and key aspects of the language. Therefore, learners had several opportunities to practice speaking English verbally. The course aimed to introduce natural English and provide students with strategies that they could apply outside of the classroom. Each session lasted 90 minutes and was held two times a week. The institute's term lasted 18 sessions, which equated to two and a half months for the study.

3.4. Instruments

The researchers used a variety of sources to get the necessary data. The primary source of data for this study consisted of the B1 Preliminary English Test (PET), a speaking pretest and posttest, a questionnaire, and an interview. The tests' reliability was

assessed by explicitly restricting the speaking themes and the amount of time that students needed to do so (Brown, 2015). Further, two experts in TEFL acted as raters and completed the scoring procedure. They calculated the Pearson correlation coefficient to measure inter-rater reliability and validate the correctness of the rating procedure. They also examined and established the validity of the content of the instruments. The upcoming sections provide further details on the instruments.

3.4.1. Language proficiency test

To ascertain the participants' general proficiency and to homogenize them in this respect, a sample of the Cambridge Test PET (Preliminary English Test B1) was administered. There were parts for speaking, listening, reading, and writing on this test. Level B1 was primarily designed for intermediate individuals aged 17 to 20 (Riera et al., 2024). The PET test's internal consistency for learners in this study was found to be 0.77, and it has the necessary validity as a standardized test (Nourdad & Mohammadi, 2022). Following the PET recommendations, the test duration was set to 140 minutes, and students who scored between 140 and 159 were classified as intermediate level.

3.4.2. Pretest and posttest of speaking

The researchers measured the students' speaking proficiency by giving them a speaking pretest and posttest of the TOEIC (Test of English for International Communication). The test was selected mainly because it was a pre-made, standardized test with a special scoring scale. There were eleven items on the test, ranging in difficulty from easy to challenging. The test guideline outlined the test's basic structure and indicated which speaking skill components would be assessed by each question. The tasks in each section include two questions about reading a text aloud, one about describing a photo, three questions about answering questions, three questions about responding using provided information, one question about offering a solution, and one about expressing an opinion.

The posttest was created by rearranging the pretest elements. According to Ordem (2023), all spoken languages are transient. It is impossible to repeat what someone says at a certain location and time, although it can be recorded. Because speech is spontaneous, context-dependent, and dynamic, learners are more likely to make mistakes in a natural situation. Each speaking test, whether pretest or posttest, took thirty minutes to be completed. The values of Cronbach's Alpha estimated for the pretest and post-tense were 0.74 and 0.88, respectively. It indicated that the reliability values for the pretest scores were 'acceptable' and the reliability values of the posttest were 'very good' values based on the reliability standards suggested by DeVellis (2021).

3.4.3. TOEIC speaking test scoring scale

Following the standardized guidelines in the Examinee handbook: TOEIC Speaking & Writing tests (2022), Question 1 and Question 2 assessed pronunciation, while Question 3 assessed vocabulary and grammar. Questions 4, 5, and 6 evaluated cohesiveness in addition to grammar, vocabulary, and pronunciation. The procedure was repeated until all components in Questions 10 and 11 were assessed. Two raters were requested to take part in the scoring procedure. It is worth noting that both evaluators (one female and one male), one aged 48, and the other aged 55, possessed Ph.D. in TEFL and had over twenty years of teaching experience focusing on education for language teachers and English language assessment. Following the guidelines, item 11 was scored on a scale of 0–5, while items 1–10 were assessed on a scale of 0–3. Ultimately, the evaluations were added together and transformed into a number between 0 and 200.

3.4.4. Questionnaire

The attitude questionnaire, which was created based on existing literature, was used to gauge how the participants felt about the Photovoice method. The researcher followed Dornyei and Taguchi (2009) in constructing, administering, and processing the questionnaire. About 12 five-point Likert scale items that varied from strongly agree (5) to strongly disagree (1), made up the questionnaire. The Participants were asked to rate how Photovoice affected their self-esteem, motivation, ability to overcome negative emotions while speaking, autonomy, desire to speak in English, interactions with teachers and other students, creative expression, managing fear of speaking, and overcoming lack of confidence issues. The questionnaire was distributed manually, and only those who were in the experimental group responded within ten minutes. According to the reliability criteria proposed by DeVellis (2021), the questionnaire's reliability was 0.85, which was regarded as a 'very good' score. Moreover, two experienced EFL experts validated the questionnaire's content, confirming its relevance.

3.4.5. Interview

The researchers adopted Sa'adah et al.'s (2023) perspective on interviews as tools for engaging participants and exploring their viewpoints, while the interview framework consisting of four questions was developed following Dörnyei and Taguchi's (2009) guidelines. Thirty students in the treatment group participated in a semi-structured interview to learn more about their attitudes and discuss their thoughts on using Photovoice in their speaking sessions. They were questioned about the advantages and disadvantages of employing the Photovoice method along with the ideas they had to increase the impact of learning through this method. They were also asked if they liked using Photovoice for their speech practice. The interview proceeded with a request for participants to express their thoughts on the Photovoice method's overall efficacy in teaching English. Each interview,

conducted in English, lasted approximately 20 to 25 minutes. The intra-rater reliability of the coding made for the common themes of the interview amounted to 95%.

3.5. Data collection procedure

A total of 60 female intermediate EFL learners were selected through convenience sampling from a pool of 150 language learners enrolled at a private English institute in Ardestan, Iran. Based on their performance on the PET test, they were homogenized, with thirty placed in the experimental group and thirty in the control group. Before receiving the specific treatment involving Photovoice, participants took a TOEIC speaking pretest to assess their basic speaking skill. The intermediate Top Notch 1B served as the primary textbook for both groups. The experimental group engaged in the Photovoice method, expressing their viewpoints through photography and storytelling. Meanwhile, the control group focused on conventional speaking skill development methods, including role-playing, group discussions, and repeated drills.

In general, the treatment group experienced the Photovoice method along with the CPS model. In the first stage, the instructor introduced and implemented the Photovoice method. The students were given a speaking assignment or issue to explore, helping them understand related challenges. Next, they were divided into small groups to discuss the pictures and give presentations. They could work individually or collaboratively to generate ideas and explore different responses or approaches to the speaking activity. Students strengthened their speaking skills by producing images that reflected their viewpoints, sometimes capturing pictures that highlighted specific issues.

The second stage involved the practical application of the instructions. Students organized their thoughts, created outlines, and shaped their presentations. Teachers provided comments to help generate ideas and ensure alignment with the speaking exercise objectives. An LCD projector display was used to enhance students' attention and support the lecture. For each image, students responded to the 3**W**s: **W**hat was taking place in the picture, **W**hy the picture was taken, and **W**hat could be done about it. Teachers also organized speaking engagements or rehearsal presentations where students received constructive feedback from both teachers and peers.

Lastly, students were encouraged to analyze themselves at the reflecting and evaluating stage. They reviewed and compared images during sessions, allowing them to assess beneficial adjustments for overcoming obstacles and improving their speaking skill. The teacher conducted a qualitative analysis of students' descriptions to explore their educational experiences. A schematic representation of the three phases is summarized in Figure 1.

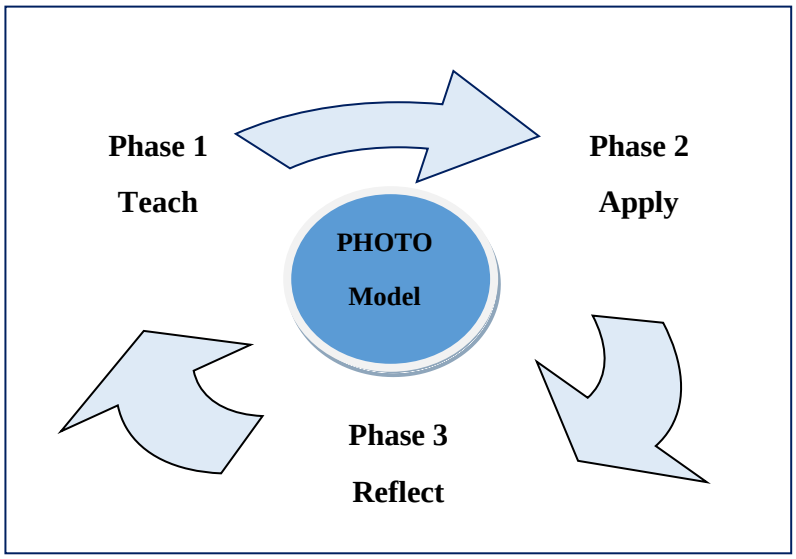

Figure 1. A schematic representation of the study's PHOTO model

At the end of the semester, photographs were displayed on the walls of the institute's spacious hall. Participants stood next to their photos and presented them in a gallery walk to other students and staff. The gallery walk allowed students to showcase their individual or group projects on posters while moving around to observe and discuss other photos, similar to how painters exhibit their artwork. Various questions were posed to gather further information, elaborations, and explanations on the issues under discussion. The procedure provided an opportunity to enhance their English-speaking skills and voice their concerns and needs in front of an audience.

During the procedure, thirty students from the experimental group responded to an attitude questionnaire, and all were interviewed. For a more thorough thematic content analysis, the interviews were audio recorded and then transcribed. Participants were informed about the recordings, and their consent was obtained at every stage of the data collection process. The researchers

aimed to identify primary and recurring themes. To maintain confidentiality, the researchers used codes or pseudonyms instead of real names when quoting participants' comments.

Nonetheless, the control group continued with conventional speaking exercises, including role-playing, subject discussions, and debates. Each session began with a few questions about the previous lesson. The instruction followed three main phases: grammar, logic, and rhetoric. Grammar focused on fundamental rules in the early stages, logic emphasized reasoning in intermediate levels, and rhetoric aimed at persuasive speaking and clear expression in advanced stages. Teacher-led activities included reading, explanations, and discussions, with no visual methods like Photovoice. After the instruction and treatment sessions, both groups took the speaking posttest. Once all data were collected, the analysis was conducted based on the study's objectives and data type.

3.6. Data analysis procedure

The researchers used SPSS, version 26 to analyze the data. The analytical phase included summarizing the pretest and posttest data and applying descriptive and inferential statistical methods, such as paired samples t-tests and one-way ANCOVA. The two-group pretest/posttest design, which compared the effects of two interventions using before-and-after measurements, required ANCOVA (Pallant, 2020). Since these groups could have differed in multiple ways, ANCOVA helped minimize some discrepancies. It was also useful when researchers relied on pre-existing groups rather than randomly assigning participants (Pallant, 2020). Additionally, descriptive statistics for the attitude questionnaire and thematic content analysis of the interviews were used to address the second research question. Finally, a cross-check was conducted, comparing data from test scores, participants' attitudes, and interviews.

4. Results

B1 Preliminary English Test (PET) was run to determine the participants' proficiency level before the study. The results of the descriptive statistics for the PET are given in Table 2.

Table 2. Descriptive statistics for the PET

N Valid	60
Mean	148.75
Median	147.00
Mode	147.00
Std. Deviation	4.36
Variance	19.03
Skewness	.25
Std. Error of Skewness	.30
Kurtosis	-.71
Std. Error of Kurtosis	.60
Range	19.00
Minimum	140.00
Maximum	159.00
Sum	8925.00

The mean score represented the average PET score across the sample, while the median and mode indicated the midpoint and the most frequently occurring scores, respectively. The standard deviation showed that the scores deviated from the mean by approximately 4.36 points, with a variance of 19.03. The distribution exhibited a slight positive skew, and the kurtosis value suggested it was somewhat flatter than normal. The PET scores ranged from 140.00 to 159.00, with a total range of 19.00. The sum of all PET scores was 8925.00.

4.1. Results of pre/post tests

The researchers examined two assumptions before performing an ANCOVA: first, linearity to determine whether a linear relationship exists between the dependent variable (posttest) and the covariate (pretest), and second, regression slope homogeneity. The scatter plots for each group indicated that the linearity assumption was not violated, as they showed a linear (straight-line) relationship, rather than a curved one (See Figure 2).

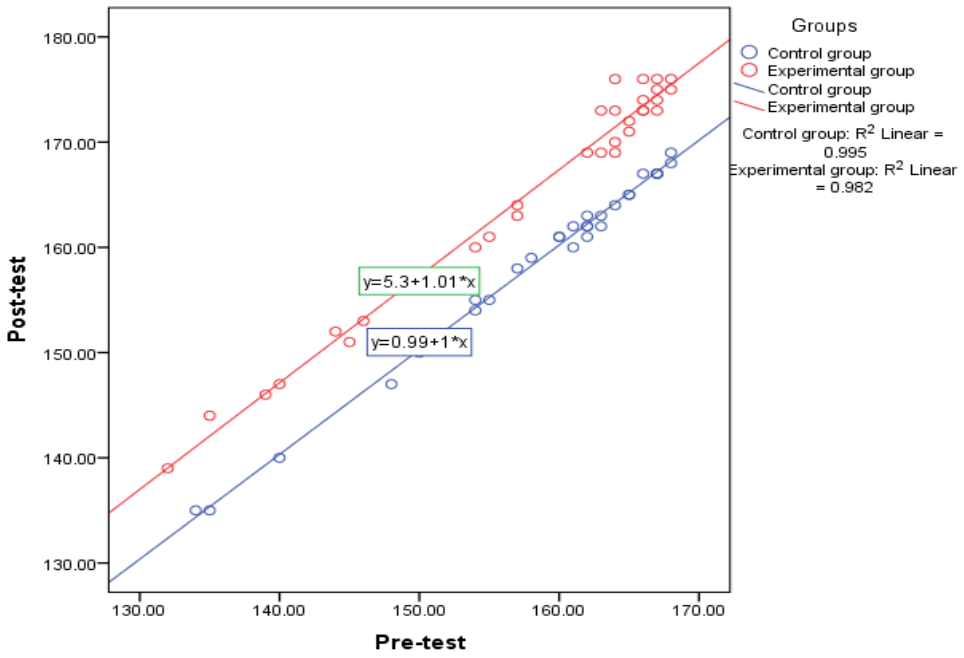

Figure 2. Scatter plot

It was evident from the research design that the groups were randomly selected and the samples were independent. The absence of interaction between the covariate and posttest scores supported the assumption of consistent regression slopes. Trimmed means and skewness analysis were used to explore the normality assumption. The results are presented in Table 3.

Table 3. Test of normality assumption

			CG	EC
Pretest	Mean		144.32	144.31
	95% Confidence Interval for Mean	Lower Bound	143.62	143.60
		Upper Bound	145.02	145.01
	5% Trimmed Mean		144.37	144.33
	Skewness		-.43	-.08
	Kurtosis		-.38	-.44
Posttest	Mean		144.71	150.11
	95% Confidence Interval for Mean	Lower Bound	144.01	149.45
		Upper Bound	145.41	150.77
	5% Trimmed Mean		144.74	150.19
	Skewness		-.46	-.62
	Kurtosis		-.52	-.30

The skewness and kurtosis values, as presented in Table 3, were within the acceptable range of ±2, indicating that the data met the normality assumption proposed by Tabachnick et al. (2013). Additionally, the trimmed means fell within the lower and upper bounds of the 95% confidence interval for the mean. Therefore, the analysis confirmed that the pretest scores followed a normal distribution. Levene's F test was conducted to assess the homogeneity of variances among the groups, with the results presented in Table 4.

Table 4. Homogeneity of variances test

	Levene's Test for Equality of Variances F	Df	Sig.
Pretest	.01	1	.88
Posttest	.04	1	.84

The data presented in Table 4 showed that the variances of the two independent groups were nearly the same (p>.05). Likewise, the Sig. Values in this case were much higher than 0.05, indicating that the variances were homogenous or that the assumption of equal variances was not broken (Pallant, 2020). The homogeneity of regression slopes was investigated as shown in Table 5.

Table 5. Tests of between-subjects effects

Dependent Variable: Posttest

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Groups	437.40[a]	1	437.40	131.74	.00	.69
Pretest	.00	1	.00	.00	.97	.00
Groups * Pretest	.46	1	.46	.32	.57	.00
Total	1304560.37	60				
Corrected Total	629.95	59				

a. R Squared = .69 (Adjusted R Squared = .68)

Regression slope homogeneity was tested to assess the interaction between the covariate and the independent variable in predicting the dependent variable. A substantial interaction would invalidate the ANCOVA results (Pallant, 2020). The significance value for the Group*Pretest data exceeded 0.05, confirming that the assumption of regression slope homogeneity was met. After verifying the assumptions of normality, equal variances, and consistent regression slopes, descriptive statistics were calculated to summarize the test results. Table 6 shows the descriptive statistics for both groups' pre/posttest speaking scores.

Table 6. Descriptive statistics for the pretest and posttest scores of the EG and CG

Tests	Groups	N	Mean	Std. Deviation
Pretest	CG	30	144.32	1.87
	EG	30	144.31	1.88
Posttest	CG	30	144.71	1.87
	EG	30	150.11	1.76

The data in Table 6 indicated that, before the intervention, the speaking skill of both groups were nearly identical. This consistency raised the possibility that any variations in the posttest results were due to the Photovoice intervention rather than innate variations. Compared to the control group, which showed just a modest rise from the pretest to the posttest, the experimental group's mean posttest score climbed significantly after the intervention. This implied that the Photovoice treatment might have contributed to enhancing the speaking skill of the group that received the intervention. Then, a one-way ANCOVA was used to investigate how the Photovoice mediations affected the speaking abilities of the students. Table 7 indicates the statistical significance of the disparity.

Table 7. ANCOVA test results (posttest)

Dependent Variable: Posttest

	Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Contrast	437.40	1	437.40	131.74	.00	.69
Error	192.55	58	1.42			

After controlling for the covariate, the ANCOVA results demonstrated that the two groups' posttest scores differed significantly. Besides, the effect magnitude met Cohen et al.'s (2018) criterion of 'strong'. A paired samples t-test was also conducted to determine whether the mean difference between the pretest and posttest scores within the groups was statistically significant. The results are provided in Table 8.

Table 8. Paired samples t-test for the pretest and the posttest

Groups			Mean	Std. Deviation	T	df	Sig. (2-tailed)
CG	Pair 1	Pre/Post Test	-.39	2.05	-1.04	29	.30
EG	Pair 1	Pre/Post Test	-5.80	1.52	-20.83	29	.00

The results indicated that the experimental group had made more improvement than the control group within the groupings. The paired samples t-test results showed that the increase was statistically significant for the experimental group exclusively (P≤.01). Therefore, ANCOVA and paired t-test results contradicted the null hypothesis and indicated that the Photovoice intervention had a statistically significant impact on students' speaking skill. Figure 3 illustrates participants' pre/posttest speaking performance.

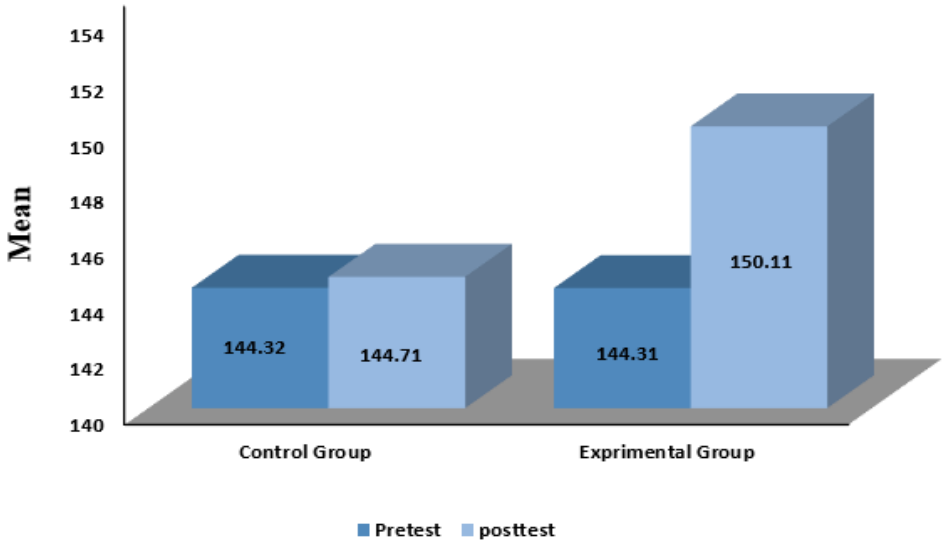

Figure 3. Groups' performance on the pretest and posttest of speaking

The bar graph depicted that participants in the experimental group outperformed those in the control group. In other words, the posttest results showed that the individuals who benefited from the Photovoice intervention outperformed the control group by a considerable amount. It was concluded from this difference that the Photovoice intervention successfully improved the experimental group's speaking skill.

4.2. Results of questionnaire

The descriptive results also provided insights into the students' attitudes toward using Photovoice to improve their speaking skill. Various aspects were examined, including self-confidence, motivation, overcoming negative emotions, independence, and creativity in speaking. These aspects helped to explain how Photovoice influenced different dimensions of the students' speaking abilities. The results are summarized in Table 9.

Table 9. Descriptive statistics for the attitude questionnaire

Statement Photovoice …	Mean	Std. Deviation	N
1. increased my confidence in communicating in English.	4.03	.80	30
2. encouraged me to use English when I speak.	4.10	.75	30
3. helped me overcome unpleasant speech-related emotions.	3.93	.73	30
4. enabled me to be self-sufficient while speaking in English.	3.86	.81	30
5. made me want to speak more in English.	4.06	.78	30
6. enabled me to engage with my peers more.	4.20	.80	30
7. allowed me to engage with my teacher more.	4.10	.71	30
8. enhanced my ability to express myself creatively.	4.33	.75	30
9. encourage me to work more cooperatively with my peers.	4.06	.73	30
10. increased my involvement in the speaking class.	4.03	.66	30
11. helped me manage anxiety when speaking in public.	3.80	.84	30
12. made it easier for me to get over my speech phobia.	4.06	.73	30

The descriptive statistics in Table 9 indicated that the Photovoice project had a positive impact on various aspects of the students' attitudes toward speaking in English. The high mean scores and levels of agreement across all statements suggested

that the students experienced enhanced self-confidence, motivation, emotional support, creativity, and interaction with peers and teachers because of participating in Photovoice activities. These findings emphasized the multifaceted benefits that were perceived by the students in their speaking skill development through Photovoice.

4.3. Results of the interview

The thematic content analysis of the interviews revealed that Photovoice was perceived as a valuable and innovative method for enhancing English learning. It stimulated students' creativity, enhanced engagement, and improved communication skills. However, technical challenges, time limitations, and subjectivity issues were weaknesses of this method. Participants suggested integrating Photovoice with traditional methods, providing clear guidelines, and enhancing technological accessibility to maximize effectiveness. Some emphasized the importance of technical support for students unfamiliar with photography equipment to minimize learning barriers. Others highlighted the need for continuous assessment strategies to address subjectivity concerns and ensure consistent evaluation criteria. Table 10 summarizes the main themes and domains reflecting participants' experiences and perceptions of Photovoice in speaking classes.

Table 10. Participants' experiences and perceptions of the photovoice in their speaking classes

Domain	Theme	Description
Strengths	1. Creativity	Foster links between visual and linguistic skills
	2. Critical Thinking	Promoted deeper exploration
	3. Engagement	Fostered a sense of involvement
	4. Visual Expression	Allowed effective expression of emotions
Weaknesses	1. Technical Difficulties	Challenges with photography equipment
	2. Time Constraints	Time required for editing and selecting photos
	3. Subjectivity	Need for standardized interpretation
Suggestions for Improvement	1. Clear Guidelines	Recommendations for clearer guidelines
	2. Technological Accessibility	Offering equipment, training, and resources to tackle technical challenges
	3. Collaborative Activities	Group exercises, debates, and role-playing
Overall Perceptions and Effectiveness	1. Enhancing Speaking Skill	Enhancing speaking, and self-expression
	2. Need for Complementary Methods and Structured Curriculums.	Combining it with other methods and integrating it into a structured curriculum with clear objectives
	3. Importance of Teacher Training	Highlighting the need for teacher training programs

In general, Iranian intermediate EFL learners' attitudes toward using Photovoice in their speaking classes were positive and they highly recognized its potential benefits while also looking for improvements and additional support to make the learning experience more successful. Many interviewees stated that their Photovoice experience helped them become more proficient speakers. They spoke English with greater assurance.

5. Discussion

The study was an effort to offer a thorough examination concerning the impact of employing the Photovoice method integrated with CPS on Iranian intermediate EFL learners' speaking skill and attitudes. The results of the statistical analysis of the speaking skill pretest and posttest showed that employing Photovoice integrated with CPS had a statistically significant impact on Iranian intermediate EFL learners' speaking skill. Moreover, the findings from the interviews and attitude questionnaires provided information about the Iranian intermediate EFL learners' attitudes toward using Photovoice in their speaking classes. Overall, the learners showed highly positive attitudes toward this method.

Reviewing pertinent literature showed that the study's results are in agreement with the findings of several researchers (Ferdiansyah et al., 2020; Mudra et al., 2023; Boamah et al., 2022). The finding revealed that it is beneficial to employ the photovoice method in language classes because it gives students additional opportunities to learn new words, expand their vocabulary, and work on their pronunciation. Photovoice was proven to enhance student productivity in speaking classes, help students become self-sufficient learners, provide an engaging learning environment, and boost students' self-esteem, which was also supported by Putri (2018). Using their imaginations, the students could explain each picture and explain what happened. In this regard, Mudra et al. (2023) claimed that Photovoice was successful in assisting people to communicate negative emotions. As a result, research indicated that the usage of Photovoice somewhat compensated for learners' lack of exposure to the target language and enhanced their language learning experiences.

From the attitude questionnaire, it was found that Photovoice made the speaking class more productive for the learners, aided them to become independent learners offered a motivating way for them to learn, and increased their self-esteem, which is also supported by Boamah et al. (2022). The learners had the choice to describe each photo by expressing their ideas about what happened. This media also enabled learners to talk about what frustrated them while speaking in front of their classmates. In this

regard, Sa'adah et al. (2023) asserted that Photovoice was effective in helping individuals to express their negative feelings. Hence, it appeared that the use of Photovoice partly compensated for the lack of learners' exposure to the target language and maximized their language learning experiences.

The data from the interview also revealed that even the shy learners in the treatment groups were satisfied with the Photovoice and were actively engaged in the class activities. Some studies indicated that the learners had highly positive attitudes toward employing Photovoice in speaking classes (Afifah et al., 2023; Ordem, 2023; Putri, 2018). In line with the findings of Ordem's (2023) study, Photovoice enhanced the learners' overall speaking abilities by supporting them in overcoming negative emotions while speaking in English. Additionally, consistent with Purti's (2018) research, the study found that Photovoice was inspiring since it gave the students a unique means of expressing themselves in English. The students' greater engagement in class, including their readiness to speak in front of the class and take part in debates, was indicative of this motivating factor (Afifah et al., 2023).

6. Conclusion

Employing Photovoice integrated with CPS in the teaching-learning process had several advantages, such as developing cooperation abilities, producing effective visual materials, encouraging creativity, and raising self-esteem. Students benefitted from Photovoice not just in terms of effectiveness but also in terms of increased motivation and interest in speaking. They appeared to like taking pictures, and they were also able to express their creativity through the camera. Stated differently, the participants' images allowed them to interpret their surroundings and represent the positive and negative aspects of their communities. They could also improve their ability to think critically and communicate ideas through the presentation. Their presentations demonstrated their capacity to explain or visualize the image. The results showed that the participants' speaking skills were positively influenced by the use of Photovoice.

To gain a deeper understanding of these findings, it is important to understand how Photovoice and CPS facilitated these changes to occur. Photovoice allowed the students to speak more meaningfully and authentically because they were able to personalize their photographs concerning their daily lives (Nisa, 2021; Sa'adah et al., 2023). This personalization aspect may have caused them to speak more and given them power. Moreover, the CPS model also enhanced students' creative thinking in solving and analyzing problems, which may lead to their critical thinking ability and speaking performance. These results are aligned with previous research (e.g., Bhatti, 2021; Ferdiansyah et al., 2020; Sa'adah et al., 2023), where interactive and visual methods have been shown to support deeper learning and engagement in language learning.

The findings also indicated that Photovoice promoted the application of more inventive and dynamic language techniques, including improved critical thinking, contextual word usage, and story building. Utilizing Photovoice integrated with CPS model not only made language production easier but also provided the learners with more interesting and meaningful learning experiences. Through the use of visual prompts, students were able to ground their language output in real-world situations, resulting in a more genuine and authentic context for teaching and learning the English language (Ordem, 2023). Moreover, most students reported feeling more motivated, self-assured, and satisfied with their language learning process.

Overall, Photovoice increased engagement and confidence in EFL learners by creating a more creative, dynamic, and supportive learning environment that solves frequent obstacles. As EFL education develops, integrating such cutting-edge methods can be crucial to generating more student-centered learning experiences, which eventually prepare students for communication in the real world. The utilization of Photovoice in education shifts the focus from community development to learning objectives and skill development. The dynamic activities that Photovoice participants engaged in had a significant impact on how they constructed sentences. Additionally, it provides instructors with deeper insights into students' self-transformation during ELT learning while also helping them understand the development of students' ELT knowledge and skills. Therefore, this teaching method may be significantly applicable to schools, universities, and other English learning environments that require specialized skills.

The research findings may also have some implications and applications for language learners and teachers, future researchers interested in English speaking skills, and curriculum designers. Students benefit greatly from photovoice, which stimulates their interest and inspires them to communicate in English. Peer interaction through the Photovoice method allows students to utilize English in actual situations, which makes them more fluent and confident. Using participatory methods like Photovoice in TEFL can also help teachers make language learning more convenient. Policymakers can help teachers overcome the challenges of teaching speaking skill by supporting and guiding curriculum developers to incorporate engaging content in the curriculum. Utilizing Photovoice combined with the CPS model at the macro level can change the learning process into a learner-centered and problem-solving-oriented one.

Although the application of the Photovoice methodology was extremely useful, it also had some limitations. The most significant limitation of the study was that because numerous commentaries can be presented on the photographs, their interpretations were relatively difficult for the researchers considering personal judgments and biases in the results. Because Photovoice relied heavily on experience and viewpoint, what was learned could be typical of specific traits among the participants rather than general trends in a larger population. Taking the students' viewpoints and researchers' biases into consideration could lead to another description of the events that were recorded.

Further studies might be oriented towards creating standard parameters or models of photo interpretation to reduce biases and ensure that meanings intended are suitably documented. Such policies need to include methods through which

participants verify their correctness in interpretation to refrain from the use of researchers' own interpretations. Future studies could also address these limitations by investigating the effects of changing the number of photos and evaluating how this influences participant engagement and the overall quality of data collection. More importantly, even if every attempt was made to evaluate and interpret qualitative data objectively, bias could have inevitably found its way into the findings. Thus, the results should not be applied generally to all situations. Lastly, it is crucial to highlight that the researchers fully accepted responsibility for any potential mistakes and flaws in the study.

7. References

Afifah, W. (2023). Capturing moments to analyze students' transformative thinking through Photovoice in an English language teaching classroom: Students' transformative thinking in ELT. *English Language Teaching Educational Journal*, *6*(3), 172-187. https://doi.org/10.12928/eltej.v6i3.10077

Ashimovna, M. N. (2022). Characteristics of successful speaking activity in ESP classes. *Ta'lim Fidoyilari*, *5*(9), 277-281. https://doi.org/10.5281/zenodo.6581878

Banitalebi, Z., & Ghiasvand, F. (2023). The representation of learning-oriented assessment (LOA) practice in EFL contexts: A case study of teacher questioning strategies. *Language Related Research*, *14*(3), 69-95. https://doi.org/10.29252/LRR.14.3.3

Bhatti, M. S. (2021). Teaching speaking skills through role play at Elementary level: An analysis. *Jurnal Arbitrer*, *8*(1), 93-100. https://doi.org/10.25077/ar.8.1.93-100.2021

Boamah, S. A., Yous, M.-L., Weldrick, R., Havaei, F., & Ganann, R. (2022). Using Photovoice as a method for capturing the lived experiences of caregivers during COVID-19: A methodological insight. *International Journal of Qualitative Methods*, *21*, 1-11. https://doi.org/10.1177/16094069221137494

Brown, H. D. (2015). *Teaching by principles: Language assessment (4th ed.)*. Pearson Education.

Ciolan, L., & Manasia, L. (2024). Picturing innovation in higher education: A Photovoice study of innovative pedagogies. *Active Learning in Higher Education*, *00*(0), 1-31. https://doi.org/10.1177/14697874241245350

Cohen, L., Manion, L., & Morrison, K. (2018). *Research methods in education (8th ed.)*. Routledge.

DeVellis, R. F. (2021). *Scale development: Theory and applications (5th ed.)*. Sage Publications.

Dörnyei, Z., & Taguchi, T. (2009). *Questionnaires in second language research: Construction, administration, and processing*. Routledge.

Educational Testing Service. (2022). *Examinee handbook: TOEIC speaking & writing tests.* Educational Testing Service. Retrieved from https://www.ets.org

Ferdiansyah, S., Widodo, H. P., & Elyas, T. (2020). Photovoice in the English as an additional language (EAL) writing classroom: No need to rush to love writing because love will grow with time. *Journal of Asia TEFL*, *17*(1), 269-279. https://doi.org/10.18823/asiatefl.2020.17.1.19.269

Ghoushchi, S., Yazdani, H., Dowlatabadi, H., & Ahmadian, M. (2021). A multimodal discourse analysis of pictures in ELT textbooks: Modes of communication in focus. *Jordan Journal of Modern Languages and Literatures*, *13*(4), 623-644. https://doi.org/10.47012/jjmll.13.4.2

Graziano, K. J. (2004). Oppression and resiliency in a post-apartheid South Africa: Unheard voices of Black gay men and lesbians. *Cultural Diversity and Ethnic Minority Psychology*, *10*(3), 302-316. https://doi.org/10.1037/1099-9809.10.3.302

Karimi, M., Chalak, A., & Heidari Tabrizi, H. (2018). Employing participatory learning through the use of photovoice in Iranian EFL learning settings: The case of reflective narratives. *Journal of Teaching Language Skills (JTLS)*, *37*(2), 65-90. https://doi.org/10.22099/jtls.2019.30359.2556

Karimi, M., Chalak, A., & Heidari Tabrizi, H. (2019). Employing Photovoice to raise intercultural sensitivity and reflective thinking of male and female Iranian EFL learners. *Journal of English Language Pedagogy and Practice*, *12*(24), 91-114. https://doi.org/10.30495/JAL.2019.671928

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development.* Prentice-Hall.

Koltz, R. L., Odegard, M. A., Provost, K. B., Smith, T., & Kleist, D. (2010). Picture perfect: Using photo-voice to explore four doctoral students' comprehensive examination experiences. *Journal of Creativity in Mental Health*, *5*(4), 389-411. https://doi.org/10.1080/15401383.2010.527797

Mitchell, C., De Lange, N., & Moletsane, R. (2018). *Participatory visual methodologies: Social change, community and policy*. Sage. https://doi.org/10.4135/9781526416117

Mudra, H., Sulistiyo, U., Abrar, M., & Wulandari, B. A. (2023). Rural college EFL students' barriers in writing and publishing a research paper: A Photovoice phenomenological approach. *PPSDP International Journal of Education*, *2*(2), 203-223. https://doi.org/10.59175/pijed.v2i2.123

Nisa, E. K. (2021). Photovoice activities to teach writing for high school students. *Borneo Educational Journal (Borju)*, *3*(1), 8-17. https://doi.org/ 10.24903/bej.v3i1.706

Nourdad, N., & Mohammadi, R. (2022). The impact of diary keeping on Iranian EFL learners' writing accuracy: Past tense use across proficiency levels. *Journal of English Language Pedagogy and Practice*, *15*(30), 151-170. https://doi.org/10.30495/jal.2022.704102

Ordem, E. (2023). Black lives matter in an EFL speaking class. *ELT Journal*, *77*(2), 176-185. https://doi.org/10.1093/elt/ccac048

Pallant, J. (2020). *SPSS survival manual: A step by step guide to data analysis using IBM SPSS (7th ed.)*. Routledge.

Pournia, S., Chalak, A., & Heidari Tabrizi, H. (2025). Utilizing Photovoice to enhance Iranian intermediate EFL learners' productive skills. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *44*(2), 119-141. https://doi.org/10.22099/tesl.2025.51761.3357

Putri, G. A. (2018). Photovoice implementation in teaching English guiding at English study program college of foreign language (STIBA) Saraswati Denpasar. *English Focus: Journal of English Language Education*, *1*(2), 86-95. https://doi.org/10.24905/efj.v1i2.35

Riera, M. A. L., Tubón, I. M. R., Quevedo, V. A. C., & Riera, N. P. L. (2024). Game-based learning to develop English skills among bachelor students. *Mikarimin. Revista Científica Multidisciplinaria*, *10*(1), 81-100. https://doi.org/10.61154/mrcm.v10i1.3328

Sa'adah, L., Anjarwati, R., Romadhoni, M., & Rabbianty, E. N. (2023). Students' voices of teachers' online written feedback. *OKARA: Jurnal Bahasa dan Sastra*, *17*(1), 118-133. https://doi.org/10.19105/ojbs.v17i1.8441

Sarani, A., Zarei, M. J., & Navidinia, H. (2020). Effect of online flipped classroom on students' writing development at senior high school. *Journal of English Language Teaching and Learning*, *12*(26), 495-515. https://doi.org/10.22034/elt.2021.44600.2348

Saslow, J., & Ascher, A. (2015). *Top Notch: English for today's world. Top Notch (3rd ed.)*. Pearson Education.

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2013). *Using multivariate statistics (6th ed.)*. Pearson Boston.

Treffinger, D. J., Isaksen, S. G., & Stead-Dorval, K. B. (2023). *Creative problem solving: An introduction*. Routledge. https://doi.org/10.4324/9781003419327

Ur, P. (2021). *A course in English language teaching (2nd ed.)*. Cambridge University Press. https://doi.org/10.1017/9781009024518

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

Wang, C., & Burris, M. A. (1994). Empowerment through photo novella: Portraits of participation. *Health Education Quarterly*, *21*(2), 171-186. https://doi.org/10.1177/109019819402100204

Wass, R., Anderson, V., Rabello, R., Golding, C., Rangi, A., & Eteuati, E. (2020). Photovoice as a research method for higher education research. *Higher Education Research & Development*, *39*(4), 834-850. https://doi.org/10.1080/07294360.2019.1692791

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 87-107

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.8828.1083

The Neuroimaging Aspects of Second Language Learning in Educational Environment: A Systematic Review of Literature

Mehran Sayyadi[1]
Reza Khany[2*]
Mehrak Rahimi[3]

[1]Ph.D. Candidate in Applied Linguistics, Faculty of Literature and Humanities, Department of English, Ilam University, Ilam, Iran
[*2]Professor of Applied Linguistics, Faculty of Literature and Humanities, Department of English, Ilam University, Ilam, Iran
[3]Associate Professor of Applied Linguistics, Faculty of Humanities, Department of English, Shahid Rajaee Teacher Training University, Tehran, Iran

ARTICLE HISTORY

Received: 15 December 2024
Revised: 18 January 2025
Accepted: 08 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: r.khany@ilam.ac.ir

ABSTRACT

Neuroimaging has a substantial pedagogical advantage when applied to language learners who are experiencing cognitive procedures during educational settings. Among such learners are L2 learners who may encounter new experiences that neuroimaging may be able to clarify. The purpose of this study is to investigate whether neuroimaging tools are capable of revealing more information about the experience of L2 learners. A systematic review of the literature on the use of neuroimaging in the context of L2 learning is presented in the study. According to a comprehensive search of numerous known databases, 19 articles met the criteria for extensive analysis. The results of the analysis indicate that neuroimaging tools can be used to reveal brain function during L2 learning. Furthermore, it may enable us to uncover brain function during the learning of L2 skills and its neural responses to them. Researchers have also mentioned the challenges in the implementation of neuroimaging tools in L2 procedures. In addition to presenting the testimonies found in the literature, it is emphasized that neuroimaging in L2 learning remains relatively unknown.

KEYWORDS: Systematic review; English; L2; Neuroimage

1. Introduction

The ability to communicate in a second language (L2) is crucial for success in today's globalized world. As the process of learning an L2 significantly impacts both the functional and structural aspects of the brain (Mönte et al., 2002), neurolinguistics scholars focus on investigating brain functions to understand language processing more deeply (Alduais et al., 2023). It is now believed that language learning involves multiple regions of the brain, challenging the conventional view that language is primarily a left-sided cognitive function (Lindell, 2006; Quin-Conroy et al., 2024). Drawing on the fact that L2 learning involves many areas of the brain and induces changes in brain functions (Osterhout et al., 2008) intrigues necessity to understand the temporal dynamics, underlying mechanisms, and specific brain regions for defining the concept of L2 learning. Despite extensive research on L2 learning, the specific nature, timing, and manifestation of brain changes during this process remain largely unknown.

Tracking brain changes scientifically during language learning involves utilizing various techniques including neuroimaging, cognitive testing, and physiological experimentations. In this scheme, the recent surge of interest in neuroimaging aligned with the advancement of the devices and methods of data gathering and analysis has provided insights into brain structure, function, and connectivity during language acquisition. Neuroimaging as an objective, non-invasive method for evaluating human brain function, employs quantitative computational techniques to explore brain structure and function (Fedorenko & Kanwisher, 2009). Neuroimaging aids language educationists to understand which brain areas are involved in learning and using a language,

Cite this article: Sayyadi, M., Khany, R. and Rahimi, M. (2025). The neuroimaging aspects of second language learning in educational environment: A systematic review of literature. *Applied Linguistics Inquiry*, *3*(1), 87-107. doi: 10.22077/ali.2025.8828.1083

how bilingual and monolingual brains function, how brain adapts to new linguistic experiences and challenges, and how language disorders can be detected and treated (Bialystok, 2024; Zhang et al., 2024).

Although several reviews have examined the issue of neuroimaging in language education, none have specifically focused on its application in L2 learning contexts (Aldhaheri et al., 2021; Antonicelli & Rastelli, 2023; Comstock, 2024; Sulpizio et al., 2020). Developing a systematic review of existing studies in this domain is crucial to consolidate the theoretical bases of neurolinguistic field and provide a foundation for further empirical studies (Angelovska & Roehm, 2023; Gernsbacher & Kaschak, 2003). This systematic review thus aims to identify neuroimaging studies related to second language learning (L2), addressing a significant gap in the SLA and seeks answering the following questions:

1.What are the findings of neuroimaging for L2 learning?

2.What are the benefits and opportunities of using neuroimaging tools in L2 learning?

3.What are the challenges of applying neuroimaging to L2 learning studies?

2. Literature review

2.1. Neuroimaging techniques

Neuroimaging is supported by principles of neuroscience, cognitive science, and physics allowing multidisciplinary researchers track the way brain learns and processes languages. Cognitive psychology, neuroimaging, and behavioral neurology, despite their individual nuances, demonstrate overarching characteristics that underscore the integration of these disciplines (e.g., Raichle, 2009; Savoy, 2001). Neuroimaging has been a crucial component of cognitive neuroscience and mental health research during the past few decades, which has greatly enhanced our understanding of neural processes that affect cognition and behavior, as well as their modifications associated with psychiatric and neurological disorders. In recent years, computational methodologies have become increasingly pivotal in cognitive neuroimaging research (Loosen et al., 2024). By elucidating neurobiological processes and correlating experimental data with underlying mechanisms, these frameworks establish a basis for understanding complex brain-behavior interactions and predicting cognitive, behavioral, and clinical outcomes (Blanken et al., 2021; Kraus et al., 2023). Neuroimaging research has yielded numerous neurobiological insights that have transformed our understanding of learning and cognitive development (Reber, 2013; Van Atteveldt et al., 2018).

In cognitive neuroscience, various methods are employed, with the most frequently utilized in neurolinguistic studies being electroencephalography (EEG), event-related potentials (ERPs), transcranial magnetic stimulation (TMS), functional near-infrared spectroscopy (fNIRS) and Positron emission tomography (PET) (e.g., Covey et al., 2024; Kram et al., 2024; Kumar et al., 2024; Momenian et al., 2024; Provost et al., 2024).

Electroencephalography (EEG) is defined as the recording of brain activity in several parts of the nervous system and the representation of such information as a visual pattern (Sharma & Meena, 2024). The event-related potential (ERPs) denotes small voltages generated in the brain's neurons in response to particular actions or stimuli (Parviainen & Kujala, 2023). Transcranial magnetic stimulation (TMS) is a non-invasive neuroimaging technique that utilizes an electromagnetic field to induce an electric signal, thereby activating targeted brain regions (Andò et al., 2021). Functional near-infrared spectroscopy (fNIRS) quantifies changes in oxygenated (HbO2) and deoxygenated (HbR) hemoglobin levels in response to neuronal activity, facilitating the study of brain tissue metabolism (Russo & Senese, 2023). Positron emission tomography (PET) is a noninvasive clinical diagnostic technique that enables the assessment of biological functions at the molecular level. A PET scanner produces cross-sectional images generated from positron-emitting radioactive markers, known as radiopharmaceuticals, which are administered internally to the subject. By employing these brain analysis techniques, researchers can forecast specific learning outcomes or treatment consequences and thus gain more significant insights in comparison to behavioral tests.

2.2. Neuroimaging and language learning

The exploration of brain changes during the L2 learning process necessitates the utilization of electrophysiological activities. These activities in the brain can be detected non-invasively through the scalp using advanced techniques characterized by rapid advancements (Perret et al., 2024). The rapid advancements in methodologies for assessing human brain function have markedly improved L2 researchers' comprehension of brain electrophysiological changes during language learning. Further, behavioral methods such as dichotic listening are integrated with neuropsychological approaches which substantiate that language construction and comprehension predominantly occur in the brain (Guiral, 2024; Wischmann et al., 2024).

In neurolinguistics, the rapid and simultaneous classification, extraction, and integration of linguistic elements are pivotal for research in L2 learning (Godfroid & Hopp, 2023; Svaldi et al., 2024). Despite the inherent complexity of this process, children naturally assimilate it from an early age. These processes are localized in specific brain regions, particularly in the left hemisphere's language-processing areas (Monroy-Sosa et al., 2021; Rajimehr et al., 2022). Language acquisition entails both quantitative and qualitative brain transformations, reflecting enhanced language proficiency and cognitive abilities (Corrigan et al., 2022).

The neuroimaging evidence on L1-L2 association indicates that access to data in L2 requires the elimination of contradictory data in the first language (Filipović, 2022; Perkins & Zhang, 2024). Languages of a learner are always perceived as highly interconnected, and even when only one language is employed, both languages are considered initiated (Zhang et al., 2020).

A number of studies conducted in recent decades have suggested that learning circumstances may alter the pathways used by the brain in order to process new information (Friederici & Wartenburger, 2010; Sulpizio et al., 2020). While non-native learners' brain feedback is generally native-like when they are transcribing language features that are similar to their own native language, it contradicts itself at other times. During comprehension, non-native learners exhibit unusual electrophysiological responses based on linguistic features partly overlapping between L1 and L2, resulting in linguistic opposition (Grey, 2023).

As literature pinpoints, a learner's L2 competence is reflected in neural activity in specific brain areas, supporting the hypothesis that neural arrangements can be modified by particular linguistic knowledge, such as learning a new language. Neuroimaging studies have demonstrated significant neural plasticity, both on a practical and structural level, when considering language exposure to a second language (e.g., Abutalebi et al., 2013; Legault et al., 2019; Wang et al., 2020; Zhang et al., 2020). The process of language processing is greatly influenced by linguistic exposure, which results in neural changes under short-term interference. Several studies have indicated that brief exposure to language can have neural effects (e.g., Fu et al., 2024; Morgan-Short et al., 2015). However, researchers have only recently begun examining the effects of short-term language instruction in controlled exploratory settings to determine to what extent and what types of involvement with a second language can influence neural responses casually.

2.3. Review studies on neuroimaging in language education

Some systematic reviews have been conducted on the relation between neuroimaging and language learning, as outlined below. Gernsbacher and Kaschak (2003) examined neuroimaging studies related to language comprehension and discovered that language comprehension involves neural computations in various brain regions, including frontal areas in the left hemisphere and their right hemisphere counterparts. The neurophysiology of language, as explored through neuroimaging studies by Démonet et al. (2005), has revealed that clinical neuroimaging is an indispensable tool for improving the efficacy of examinations, predictions, and treatment methods for individuals with brain damage related to language. Herringshaw et al. (2016) investigated hemispheric differences in language processing and concluded that the relationship between language and the brain is complex and context-dependent, with bilateral activation observed across several different experimental conditions. Ware et al. (2021) performed a systematic review on L2 learning and neuroplasticity in aging, noting that only one study utilized neuroimaging. Their results indicate that L2 learning is associated with improvements in cognitive flexibility, self-regulation, working memory, and neural connectivity. Similarly, Deldar et al. (2020) conducted a systematic review focused on fMRI studies, investigating the relationship between language and working memory. Their findings demonstrated a clear interconnection between language and working memory, supported by the activation of limbic networks, including the basal ganglia or caudate, and several right temporal areas. Despite these review studies on language and neuroimaging issues, there remains a conspicuous absence of a systematic review specifically focusing on L2 learning through neuroimaging. The present study aims to address this gap.

3. Methodology

The current study is a ‘research-focused systematic literature’ utilizing Chong and Plonsky’s (2024) typology of literature reviews carried out with the aim of assessing the quality and range of studies done on neuroimaging aspects of L2 learning and providing new research insights in this realm. The methodology for performing the study is explained in detail in the following sections.

3.1. Search strategy and selection of the focal literature

In the current systematic review, Preferred Reporting Items for Systematic Reviews and Meta-Analyses PRISMA statement guidelines were followed (Page et al., 2021). To conduct a systematic review based on PRISMA, an exhaustive examination of all research conducted on a specific topic is undertaken to address a clearly defined research question. Then, relevant research papers are selected for analysis based on a variety of inclusion and exclusion criteria. Following the selection of the studies, a systematic analysis of the results is conducted and statistical techniques are employed to synthesize results. The PRISMA principles are outlined in a 27-item checklist and a 4-phase flow diagram. The PRISMA checklist includes points concerned with every aspect of systematic reviews and meta-analyses and allows the detailed analysis of the selected articles for review addressing topics such as the title, abstract, introduction, methods, results, and discussion (Table 1).

Table 1. PRISMA 2020 item checklist (Page et al., 2021)

Selection and topic	Item #	Features	Number of items
Title	1	Title	1
Abstract	2	Abstract	1
Introduction	3	Rationale	2
	4	Objectives	
Method	5	Eligibility criteria	11
	6	Information sources	
	7	Search strategy	
	8	Study selection	

	9	Data collection process	
	10	Data items	
	11	Study risk of bias assessment	
	12	Effect measures	
	13	Synthesis of results	
	14	Reporting bias assessment	
	15	Certainty assessment	
Results	16	Study selection	7
	17	Study characteristics	
	18	Risk of bias in studies	
	19	Results of individual studies	
	20	Results of syntheses	
	21	Reporting biases	
	22	Certainty of evidence	
Discussion	23	Discussion	1
Other information	24	Registration and protocol	4
	25	Support	
	26	Competing interests	
	27	Availability of data, code, and other materials	
			Total: 27

PRISMA diagram depicts the flow of information through the different phases of a systematic review and maps out the number of records identified, included and excluded, and the reasons for exclusions (Figure 1). The diagram along with the checklist delineates the criteria for classifying, evaluating, and determining the acceptability and eligibility of studies that are to be included within the review framework and the strategies that should be deployed to carry out the review study.

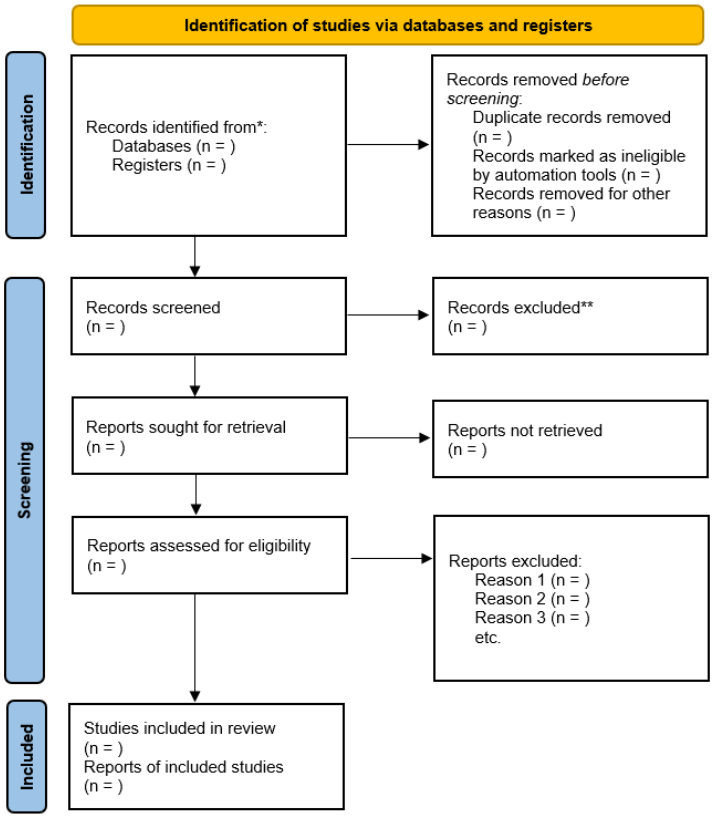

Figure 1. PRISMA 2020 flow diagram (https://www.prisma-statement.org/prisma-2020-flow-diagram)

In the pursuit of precise and accurate outcomes of the PRISMA approach, six databases were examined for this review work: Science Direct, Wiley Online Library, Sage, Taylor & Francis, Cambridge Online Library, and John Benjamin. These databases were selected due to their extensive holdings of the most frequently cited journals and articles pertinent to neuroimaging research. An extensive search was conducted in March 2024 and the inquiry was constructed using a combination of keywords in the fields of (a) Education and (b) Neuroimaging. Keywords used in the study included electroencephalography (EEG), Functional Magnetic Resonance Imaging (FMRI), Event Related Potential (ERP), transcranial Magnetic Stimulation

(TMS), Functional near-infrared spectroscopy (FNIRS), Positron Emission Tomography (PET), Magnetic resonance imaging, L2, and second language. To investigate the databases, the following terms were used: (EEG or FMRI or ERP or TMS or FNIRS or Magnetic resonance imaging or PET) AND (L2 or second language). In addition to the database query, a manual search was conducted on the following high-quality neurolinguistics journals indexed in Scopus and Clarivate: Frontiers in Human Neuroscience, Journal of Cognitive Neuroscience, Language, Cognition and Neuroscience, Bilingualism: Language and Cognition, Journal of Neurolinguistics, and Journal of Psycholinguistic Research. These journals were chosen for manual search to identify additional studies that might have been excluded due to not containing the aforementioned exact words, given their high yield in database searches.

3.2. Inclusion/exclusion criteria and search outcome

The constructed dataset comprises articles that explore the application of neuroimaging in L2 learning procedures within educational contexts. The study design, methodology, participant demographics, age, and geographic location were not restricted. However, the final selection was limited to studies conducted within L2 learning environments. Consequently, articles that investigated neuroimaging in language studies but did not specifically address L2 learning were excluded. Additionally, studies employing methods other than neuroimaging, not directly related to L2 learning, or focusing on simultaneous bilinguals rather than non-native learners from other countries were also excluded. These criteria facilitate a more profound understanding of language learning within an L2 context. L2 was defined in the study as "any language acquired after a first language system has already been established" (Littlewood, 2004, p. 502). Furthermore, neuroimaging was viewed as the method of understanding the brain through the use of spatial relationships between neurons and systems on a temporal scale (Bandettini, 2009) ranging from high (milliseconds) as in EEG devices to low (seconds) in fMRI scans to measure changes in brain activity precisely over time.

According to Figure 1, 662 articles were identified across six databases, however 212 articles were deleted due to duplication and appearance in two sets of databases. After a thorough review of the titles and abstracts of the remaining 450 articles, 248 were excluded due to their lack of relevance to L2 learning and neuroimaging directly. Additionally, 29 studies were omitted due to their non-English nature as they were written in other languages. A total of 202 articles were assessed for eligibility by reviewing their content, with 43 demonstrating no direct relevance to neuroimaging studies and utilizing associated devices to gather data so they were excluded from the review procedure. Furthermore, 53 studies were excluded for providing only general or theoretical information about L2 learning rather than specific and direct results. An additional 97 studies were excluded because their primary focus was not on L2 learning and only used L2 in their titles but there were no explanations or further details about L2 in their content so they could not be analyzed and were excluded. Consequently, 19 appropriate studies were selected for detailed analysis based on the selection process results (see Figure 2). Appendix 1 provides a summary of the reviewed studies and their characteristics.

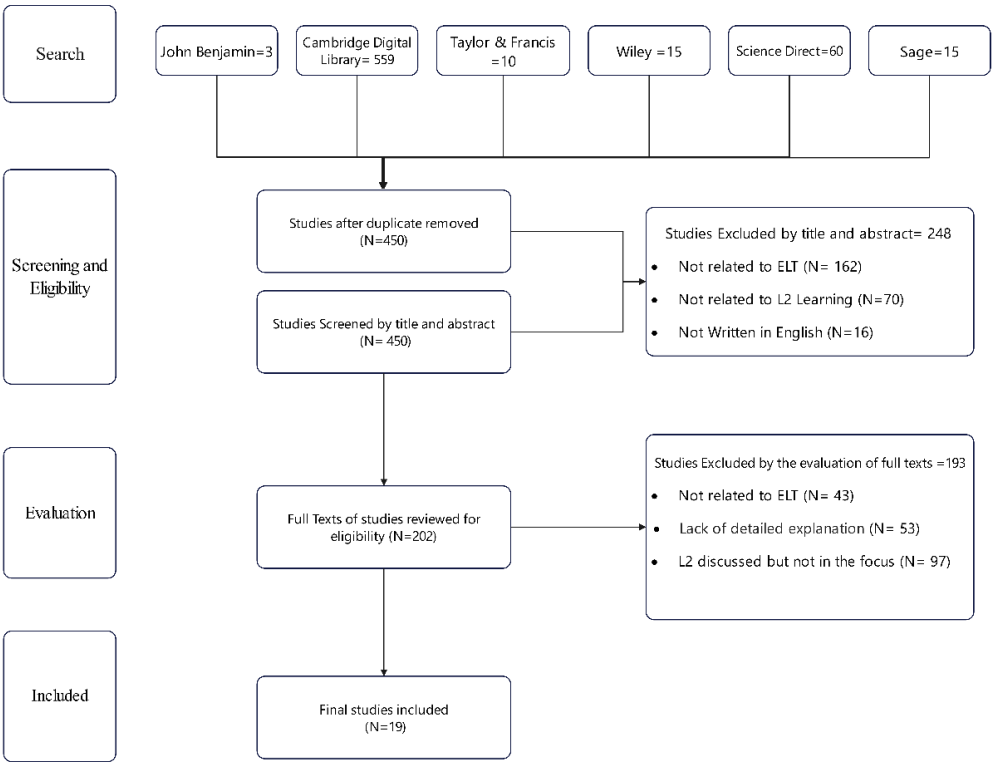

Figure 2. The diagram of study selection

3.3. Procedure

Qualitative Systematic Review Protocol was used "to minimize bias and enhance transparency and reproducibility" (Butler et al., 2016, p. 241) of this review work. As a result of conducting a careful analysis of numerous databases done both automatically and manually, evaluating the eligibility and characteristics of involved studies by following the guidelines, and integrating results by meticulous thematic analysis and integrating the results, the bias was reduced and the validity and reliability of the review procedure were increased (Shaheen et al., 2023). The following steps were taken to carry out the study (Butler et al., 2016; Davis, 2016; Mancin et al., 2024):

1) Formulating research questions to facilitate the review process: three research questions were formulated with the goal of identifying the main findings of the review studies, their contribution to L2 learning, and challenges the future researchers may encounter in this field. To be able to attain these goals, details about the design of the studies, their gathering instruments and measures, participants, type of neuroimaging devices they utilized, and types of L2 activities they used were carefully inspected.

2) Identifying a search strategy, utilizing appropriate keywords and methodologies to pinpoint pertinent literature: to be able to answer the formulated questions, a search strategy was planned and implemented by extracting the main concepts pertinent to the topic of neuroimaging aspects of L2 learning to maximize the number of relevant records retrieved in the identified databases. The following techniques were used to be able to combine suitable search terms: brainstorming keywords and phrases that describe each concept related to 'L2 education and neuroimaging' by all authors, creating a list of the keywords, selecting the keywords carefully by checking inclusion/exclusion criteria, running the primary search, evaluating the results of the first search round, and refining the keywords to come up with the final search terms (See section 3.1.).

3) Establishing review procedure pattern, involving the development of a multi-level analysis procedure that provides guidelines for study inclusion, thereby establishing a systematic framework: general requirements on how to carry out (e.g., Butler et al., 2016) and report review studies including PRISMA protocol (Page et al., 2021) as well as field-specific recommendations on how to perform secondary research in Applied Linguistics (Chong & Plonsky, 2024) were incorporated to ensure that all detailed information of the previous studies was captured and their analysis could be mapped out carefully (See section 3.2. and Figures 1 and 2).

4) Critical analysis, involving the application of methodologies to evaluate the reliability of the studies under consideration, which is essential for identifying potential biases: careful analysis of each study based on PRISMA item checklist (See Table 1) was done to ascertain that the reports incorporated the structure of a research paper and were suitable to be analyzed based on the eligibility criteria set in stages 1-3.

5) Collecting and interpreting data, involving the identification of methodologies for data collection and evaluation to ensure a comprehensive assessment of the study outcomes, by presenting different aspects of the data (Schiavo, 2019): to be able to identify and present the recurring patterns and themes in the gathered data, content analysis was carried out. The data analysis was done following the guideline suggested by Chong and Plonsky (2024) for systematic literature reviews in applied Linguistics (See section 3.4. below for details).

3.4. Data analysis

Data from the included studies in the systematic evaluation were extracted and recorded in Microsoft Excel and were subsequently categorized into two distinct sections. The first section encompasses the principal characteristics of the studies, including the author, year, country, participants, gender distribution, mean age, primary objectives, and the evaluation tools and instruments employed (Appendix 1). The second section presents the conclusions and outcomes of the evaluated studies, detailing the research objectives, and statistical measures indicating the significance of the results, along with the observed positive, neutral, or negative effects (Appendix 2). It is noteworthy to mention that the current study primarily focuses on a literature review, and the included studies did not provide effect sizes, which precluded the possibility of conducting a meta-analysis.

The coding was done in three steps based on Saldana's guidelines (2015), i.e., pre-coding, first cycle coding, and second cycle coding utilizing both manual and computer-assisted analyses.

Pre-coding phase: in this phase, the included literature was read carefully and then primarily coded by two researchers (author 1 and 3) independently.

First-cycle coding: in this phase, the primary codes established in pre-coding stage were compared in three elaboration sessions and disagreements between the coders were discussed and resolved. The inter-coder reliability was calculated by Kappa statistic and found to be .92.

The second-cycle coding: in this phase, the data were analyzed by content analysis where main themes and categories were established. Then frequencies and percentages were calculated to observe patterns and trends in the gathered data as explained in section 4. Results below. Upon interpreting the data, the findings were discussed and conclusions were drawn.

4. Results

4.1. The design of the studies and data gathering instruments

The majority of the analyzed inquiries were quantitative (89.47%) and each study employed a neuroimaging method for data collection as detailed in Appendix 1. Given the complexity of language learning, several data collection instruments were utilized to complement the neuroimaging data to either assess language proficiency at the study's onset or evaluated performance after instruction and task completion (52.63%) or examine the participants' psychological and behavioral variables (63.15%).

The tasks used in the analyzed studies were designed to measure language knowledge/competence and cognitive capabilities/variations or the combination of both in the process of neuroimaging data record. The linguistic tasks assessed lexical awareness, oral fluency, or processing of certain linguistic features. Grant et al. (2015), for instance, used lexical decision and semantic judgement tasks to examine the association between L2 learners' lexical architecture and cognitive control. Similarly, Liu et al. (2021a) utilized language switching tasks to assess the relation between language control processing and cognitive control in instructed L2 learning. Sullivan et al. (2014) incorporated verbal fluency tasks to detect neural processing modifications in the process of L2 learning during developmental stages. In the same vein, Dallas et al. (2013) used sentence processing tasks to investigate real-time processing of filler-gap dependency of L2 learners.

Psychological and behavioral measures encompass a wide range of instruments, including questionnaires and tests, used to assess participants' behavioral and psychological changes. These measures aim to support the data extracted from neuroimaging tools and provide more comprehensive results about L2 learning procedures. According to Barbeau et al. (2017), language proficiency and expertise questionnaires, as well as read-aloud protocols, were used and the results obtained with these instruments were supported by fMRI data. In their study, Grant et al. (2015) applied a language proficiency test to provide information about the participants' L2 proficiency, ensuring that any deviations in neuroimaging results caused by differences in language proficiency could be avoided, thus ensuring more reliable neuroimaging results. Alos, Sullivan et al. (2014) used vocabulary proficiency tests as pre- and post-tests to examine the relation between changes following L2 learning instruction and brain modifications throughout the continuum of bilingualism.

4.2. Participants

The research in the final dataset encompassed English learners including university (%21.05%) and school students (5.26 %) and L2 learners (73.68%). Participants of all studies were adult, with further details available in Appendix 1, although two studies (Elgort et al., 2015; Xu et al., 2019) did not specify the exact age of their participants. In all cases, the participants were L2 learners, with the exception of two works (Buchweitz et al., 2009; Xu et al., 2019) which included both native speakers and L2 learners.

4.3. Types of neuroimaging

As a result of the stimuli designed for each study in order to detect learners' brain function during language learning procedures, FMRI (52.63%) was the most frequently used neuroimaging technique for measuring brain activity. It primarily assesses brain function by measuring fluctuations in blood flow associated with neural activity, allowing researchers to identify the areas of the brain responsible for specific functions. Next comes ERP (31.57%) which is highly effective at analyzing brain activity sequences with a high degree of accuracy and precision. Furthermore, other forms of neuroimaging such as EEGs (26.31%), with the capability of detecting neuronal activity millisecond by millisecond were used. To achieve more precise results regarding language learning process, some studies used more than one type of neuroimaging technique (15.78%).

4.4. Types of L2 learning activities

Given the focus of the reviewed studies on L2 learning procedures, it is essential to identify the specific aspects of L2 learning that have been evaluated using neuroimaging tools. The majority of the studies utilized comprehension-type activities (84.21%), while the remaining used production-type learning activities (15.78%).

Comprehension-type activities encompass various tasks to concurrently or sequentially gather linguistic and neuroimaging data. In their study Nakagawa et al. (2022) utilized sentence formulation and executive function and found that the reason for sentence production difficulty in L2 is cognitive overload, which occurs during completing sentence completion activities. Koyama et al. (2014) employed visuo-spatial task processing activities and determined that the posterior lateral occipital complex displayed insufficient leftward lateralization as a consequence of increased visuo-spatial demands associated with visually complex logographic symbols, which require more right-side processing. Implementing lexical and semantic proficiency tasks, as well as cognate word processing, examined by Xu et al. (2019), indicated that distinct conjunctions have distinct modifying influences when it comes to modifying relational vagueness. Non-native speakers are more likely to experience ambiguity and, therefore, produce greater relational ambiguity as sentences are connected by conjunctions with more complex semantics than native speakers. Cognitive tasks to provide responses to upper- and lower-case sentences employed by Buchweitz et al. (2009), Choi et al. (2018), Du et al. (2023) and Koyama et al. (2014) illustrated that even non-native learners' language performance could reach the same level as native learners with training. L2 word recognition tasks based on both

semantic and lexical knowledge employed by Elgort et al., (2015), Midgley et al. (2009), and Mueller (2009) revealed that the most prevalent and effective method of learning new words is through inadvertent exposure to language usage. Nonetheless, it is essential that learners with low levels of language competence explore some supplementary, more intentional learning opportunities.

Production-type activities were assessed using neuroimaging devices while subjects were engaged in constructing sentences of semantic complexity, filling in gaps, and responding to various types of questions orally. In their study, Reiterer et al. (2009) examined the hypothesis if L2 learners in their initial phases and/or bilinguals with poor fluency and less experience rely more on right-hemisphere (RH) areas when processing their L2 using EEG signals and production tasks. Their findings were in favor of RH theories indicating that RH involvement in (late) second language learners with less experience and less training in the L2 was more evident. In another study done by Dallas et al. (2013), real-time processing of sentences containing filler-gap dependencies by late-learning ESL speakers was examined. Their results suggested that, although the L2 speakers as a group are not sensitive to plausibility variations, correlational analysis indicates that more proficient L2 speakers, like the first-language (L1) speakers, are sensitive to plausibility variations while processing filler-gap sentences. In their study, Xu et al. (2019) investigated how different conjunctions affect the interpretation of a following ambiguous pronoun for non-native speakers of English in comparison to native speakers utilizing event-related potentials (ERPs). Their findings suggested that different conjunctions exert different modulating effects on resolving referential uncertainty/ambiguity and relative to native speakers, non-native speakers are more likely to encounter referential uncertainty when the sentences are conjoined by conjunctions with more complex semantics.

4.5. Measures used

The reviewed studies provide insights into brain function and modifications during language acquisition, enabling the application of neuroimaging techniques to assess the brain activities discussed in the preceding section. Three main types of neuroimaging measures or tasks were utilized including fMRI, EEG, and ERP in the process of doing language activities.

As a neuroimaging tool, fMRI provides valuable insights into various L2-related theories. In Grant et al.'s (2015) study, which was one of the pioneering studies in neurolinguistics, lexical decision tasks were analyzed using MRI and revealed a strong relationship between processing time and task type in interpreting L1 and L2 words. Another study conducted by Yang et al. (2015) used fMRI to assess individual differences in cognitive control and lexical architecture among late L2 learners, and reported that competent L2 participants demonstrated decreased brain activity in specific brain areas after training, which demonstrated improved language learning ability. The study was conducted by Barbeau et al. (2017) examined neural changes during L2 Learning using fMRI and found that the left inferior parietal lobe was associated with improved skills in the second language. It has been demonstrated by Choi et al. (2018) that uppercase texts engage specific brain regions, while lowercase texts engage multiple language-related areas, which supports the automaticity theory of language acquisition. In a similar manner, Liu et al. (2021b) explored children's cognitive flexibility and creativity as they learned languages using fMRI. As a result of their research, it was found that long periods of classroom instruction in the second language can have significant neuroplastic effects in the brain areas responsible for language control.

In the context of L2 , ERP serves as a widely used tool for collecting neuroimaging data. Wang et al. (2007) explored neurocognitive models of lexical selection through ERPs, revealing that language switching exhibits distinct neural correlates depending on the switch direction, with no specific brain region identified as a 'language switch.' Similarly, Sulivan et al. (2014) examined neural activity changes during the developmental phases of L2 learning using ERPs, noting significant neural activity shifts following brief L2 instruction. Xu et al. (2019) analyzed native and non-native speakers' interpretations of ambiguous pronouns with complex conjunctions via ERPs, finding that non-native speakers exhibited notable referential uncertainty when handling semantically intricate structures. Additionally, Du et al. (2023) utilized both ERP and EEG to investigate whether L2-specific reading skills could mitigate the influence of native language limitations, demonstrating that adults can achieve native-like neural responses in L2 reading when trained in essential skills at optimal intensity.

As a result of practical issues and limitations in managing the data gathering sessions, EEG was not commonly used by the researchers for the collection of neuroimaging data. Using EEG techniques, Midgley et al. (2009) examined how form and meaning are perceived during word recognition, as well as the role played by semantic representations. They found that L2 learners are intensively processing form representations of non-cognate translation equivalents when they are processing printed words on a form-level level, without any facilitative interaction between form equivalents. A study conducted by Dallas et al. (2013) investigated the real-time processing of filler-gap dependency sentences by late-learning English language learners using EEG, and found that the late-learning learners did not process the filler gap sentences in the same manner as the L1 learners. An EEG study by Elgort et al. (2015) examined how accidental and situational L2 word learning affects lexical competence in the L2. Based on their findings, newly learned L2 words can be identified based on both episodic and lexical semantic knowledge, and this recognition may vary depending on the learner's lexical proficiency in L2 and the context in which the word is used.

In addition to neuroimaging data, the studies also utilized objective measures for assessing language-related variables such as language proficiency through tests (Reiterer et al., 2009; Nakagawa et al., 2022), self-report language competence questionnaires (e.g., Barbeau et al., 2017), or performance tests and tasks (Yang et al., 2015).Within the studies reviewed, the most frequently employed evaluation instruments were proficiency tests (47.05%), questionnaires (29.41%), and behavioral measures (11.76%).

5. Discussion

In this study, a review of neuroimaging studies related to L2 learning procedures is presented as well as an evaluation of how neuroimaging can be applied to research on L2 learning. There may be some physiological changes that do not have behavioral implications during neuroimaging studies, and because neuroimaging relies on behavioral patterns, it might not be possible to evaluate such changes while studying neuroimaging. Consequently, researchers should consider that language learning tasks for neuroimaging must be behaviorally structured, or other complementary instruments should be used to compensate for the lack of behavioral structures. This issue has been addressed in the majority of the reviewed articles.

Appendices 1 presents an overview of the reviewed articles' characteristics and outcome, respectively. Based on the goals of the study, 19 articles from John Benjamins, Wiley, Taylor & Francis, Sage, Science Direct, and Cambridge were evaluated. The study's findings demonstrate a notable rise in neuroimaging research within the L2 context since 2015. as a result of the growing acceptance of this method and its numerous advantages.

The outcome of the study shows that the predominant imaging technique in L2 research is functional magnetic resonance imaging (fMRI), which is likely favored due to its ability to provide detailed brain imaging. This method has been widely endorsed as an effective approach for exploring L2 learning. While there is extensive literature on fMRI, there remains a significant need for additional research utilizing event-related potentials (ERP) and electroencephalography (EEG).

5.1. What are the findings of neuroimaging for L2 learning?

A thorough examination of the literature indicates that neuroimaging is a powerful tool for investigating brain function in the context of second language (L2) acquisition. Reiterer et al. (2009) found that learners with lower proficiency exhibited greater right hemisphere involvement compared to more proficient learners. Their study also highlights the importance of proficiency level in L2 learning, as Liu et al. (2021b) demonstrated that extended L2 learning leads to significant neuroplastic changes in brain regions linked to language control, as the outcomes of other studies on L2 learning have already illustrated (e.g., Choi et al., 2018; Nakagawa et al., 2022; Wang et al., 2007). Koyama et al. (2014) discovered that reduced left-ward brain lateralization is associated with enhanced right-hemisphere spatial cognition, rather than indicating increased effort by L2 learners. These results are supported by other neurolinguistic studies, showing that the brain is lateralized for specific tasks and that language has a unique role in the brain (e.g., Friederici, 2011; Pinel & Dehaene, 2010).

Neuroimaging research has significantly advanced the study of reading and writing by elucidating brain activity during these cognitive processes (Barquero et al., 2014; Schlaggar & Church, 2009). According to Buchweitz et al. (2009), learners exhibit diverse cognitive responses and varying levels of language and cross-linguistic processing at the sentence level during writing tasks. Cattinelli et al. (2013) assert that written stimuli are crucial in reading processing's initial phases, while Desroches et al. (2010) propose that neural activity during reading offers insights into language learning mechanisms. Despite widespread interest in understanding L2 processing in the brain, researchers must proceed with caution, ensuring that their findings are rigorously validated against established research outcomes.

5.2. What are the benefits and opportunities of using neuroimaging tools in L2 learning?

The application of neuroimaging technologies in SLA research offers a wide array of advantages. Scholarly works underscore that one of the most pivotal benefits is the capability to investigate brain activity during L2 learning and assess the suitability of these tasks for improving L2 learning and their incorporation into experimental research. This assertion is corroborated by numerous studies (e.g., Rastelli, 2018; Roberts & Siyanova-Chanturia, 2013) that stress the significance of utilizing metalinguistic tasks to gain a more profound understanding of the L2 learning process and its relevance in experimental research. Additionally, neuroimaging tools provide valuable insights into the behavioral patterns of L2 learners and their alignment with brain function, aiding instructors and researchers in formulating more effective strategies for L2 learning (Buchweitz et al., 2009; Elgort et al., 2015; Reiterer et al., 2009).

Neuroimaging techniques provide a more effective method for examining cognitive processes during L2 learning compared to traditional questionnaires or tests. These tools offer direct insights into brain function during L2 learning, which is unattainable through conventional methods that lack direct access to brain functions during L2 learning (Friederici & Wartenburger, 2010). Numerous studies have corroborated the complementary function of neuroimaging in the human L2 learning process (Herholz et al., 2001; Von Rhein et al., 2015; Whelan, 2007). This capability also facilitates the exploration of brain function under diverse conditions that would be challenging or impossible to study using traditional L2 learning materials, such as analyzing sentence processing in a resting state. This broadens the scope for researchers, potentially leading to the development of more effective language learning methods (Choi et al., 2018; Lum et al., 2022; Nakagawa et al., 2022).

The integration of traditional instruments with neuroimaging studies significantly enhances the precision and reliability of findings for learners. This assertion is supported by other studies, which highlight the ability of this research approach to provide insights at an individual level (Bajracharya, & Peelle, 2023; Michon et al., 2022). Consequently, learners can develop a more comprehensive understanding of brain function in the context of language learning. However, additional research is essential to fully elucidate the complexities of language processing and learning in the human brain (Makita et al., 2013; Perpiñàn, 2015; Yongqi Gu, 2016).

Numerous academic studies have confirmed the significant role of neuroimaging in advancing L2 learning (Sabourin, 2009; Zheng, & Zhang, 2024). Beyond analyzing neural changes during L2 learning processes, neuroimaging can also uncover brain data processing during these procedures (e.g., Barbeau et al., 2017; Choi et al., 2018; Midgley et al., 2011). Additionally, neuroimaging can enhance L2 learning processes and elucidate the cognitive mechanisms involved in language learning.

The reviewed studies frequently establish a connection between the integration of L2 learning into neuroimaging research and an enhancement in the interaction among various types of language production tasks. However, most studies are unable to conclusively explain how cognitive variables may affect brain function and the duration of these benefits. Liu et al. (2021a) demonstrated that prolonged engagement in L2 learning results in neural modifications. Additionally, the systematic activation of L2, combined with various L1 writing techniques, may influence neuroimaging outcomes, as supported by other studies (Kelsen et al., 2022; Tao et al., 2021). Du et al. (2023) demonstrated that only short-term, concentrated word training could affect neural responses in L2; however, this method could not determine the long-term effects of L2 exercise. Conversely, Hu and Luo (2024) suggested the application of neuroimaging studies in L2 vocabulary retention research and its effectiveness.

Several studies have shown that the process of recapturing newly acquired L2 words may involve lexical and semantic knowledge, contingent upon the individual's L2 lexical competence and the contextual introduction of the word (Elgort et al., 2015). As indicated by Barbeau et al. (2017) and Sullivan et al. (2014), the activation of the left hemisphere of the brain enhances L2 learning and the capacity to utilize various aspects of language learning. The implementation of effective strategies aimed at augmenting brain activity during L2 learning has the potential to enhance overall brain cognitive functioning (Li et al., 2014; Osterhout et al., 2008). As a result of instruction that facilitates L2 learning, skilled L2 learners exhibit reduced brain activity in specific brain regions post-instruction. These findings indicate that instruction aimed at enhancing L2 learning can lead to improved neural function (Liu et al., 2021a; Yang et al., 2015). This advancement offers the potential to integrate neuroimaging findings into L2 instruction and research.

5.3. What are the challenges associated with applying neuroimaging to L2 learning studies?

In the context of incidental contextual L2 learning, a reliable semantic agreement effect is more pronounced in higher competence L2 learners compared to their lower competence counterparts, as supported by other studies (Bardovi-Harlig & Dörnyei, 1998; Mulder et al., 2019; Pu et al., 2024). It is also essential to recognize that L2 reading comprehension requires more time than L1 reading. This temporal disparity should be considered during data extraction in L2 reading research (Dallas et al., 2013). Given the characteristics of neuroimaging in L2 learning research, it is unsurprising that these tools have been applied in educational settings to investigate the relationship between neuroimaging and L2 learning. However, it is crucial to acknowledge that psychological changes do not always correlate with behavioral changes (Sullivan et al., 2014). Researchers must be cognizant of this issue when interpreting the data obtained through this procedure. The utilization of unfamiliar words within an instructional context leads to an elevated level of extraneous cognitive load, necessitating careful management during data manipulation (Bolkan & Goodboy, 2020; Mueller, 2009; Xu et al., 2019). It is imperative that task designers recognize that not all tasks possess discriminative capabilities, a consideration that must be factored into forthcoming neuroimaging research (e.g., Dallas et al., 2013).

6. Conclusion

The aim of this research was to investigate the use of neuroimaging in educational settings among L2 learners and to clarify the effectiveness and limitations of neuroimaging in relation to L2 learning. A systematic review, complemented by manual research, identified several relevant articles, with 19 studies meeting the inclusion criteria. The analysis showed that neuroimaging is associated with various factors, including methodological approaches, instrumentation, tasks, and brain function during L2 learning. The results suggest that neuroimaging has significant potential for investigating L2 learning compared to other methodologies.

Addressing the challenges identified in this study, particularly in task design, is crucial for enhancing the understanding of the L2 learning context in follow-up studies. Further, the exact mechanisms of brain function during L2 learning remain unclear, necessitating further research. Despite the authors' efforts to adhere to systematic review guidelines, the limitations of the current study highlight the need for a comprehensive meta-analysis on neurolinguistic studies of L2 learning. Additionally, most of the reviewed studies examined only certain aspects of the L2 learning process using neuroimaging methods, and further research on different aspects, such as integrating various skills, is highly recommended. There is also a lack of evidence regarding English L2 learners with different native languages, which future studies should address to fill this gap in the literature. Future research should incorporate neuroimaging to better understand its utility in revealing the neurological processes underlying L2 learning.

7. References

*** Denotes articles systematically reviewed.**

Abutalebi, J., Della Rosa, P. A., Ding, G., Weekes, B., Costa, A., & Green, D. W. (2013). Language proficiency modulates the engagement of cognitive control areas in multilinguals. *Cortex*, *49*(3), 905-911. https://doi.org/10.1016/j.cortex.2012.08.018

Aldhaheri, T. A., Kulkarni, S. B., & Bhise, P. R. (2021). Methods of Electroencephalography in Neurolinguistics: A Systematic Review. *Grenze International Journal of Engineering & Technology (GIJET)*, *7*(1).

Alduais, A., Alduais, A., Amidfar, M., & Alizadeh Incheh, S. (2023). Neurolinguistics: A scientometric review. *Cogent Arts & Humanities*, *10*(1), 2197341. https://doi.org/10.1080/23311983.2023.2197341

Andò, A., Vasilotta, M. L., & Zennaro, A. (2021). The modulation of emotional awareness using non-invasive brain stimulation techniques: A literature review on TMS and tDCS. *Journal of Cognitive Psychology*, *33*(8), 993-1010. https://doi.org/10.1080/20445911.2021.1954013

Angelovska, T., & Roehm, D. (2023). A Selective Review of Event-Related Potential Investigations in Second and Third Language Acquisition of Syntax. *Languages*, *8*(1), 90. https://doi.org/10.3390/languages8010090

Antonicelli, G., & Rastelli, S. (2023). Event-related potentials in the study of L2 sentence processing: A scoping review of the decade 2010-2020. *Language Acquisition*, *30*(2), 163-200. https://doi.org/10.1080/10489223.2022.2141633

Bajracharya, A., & Peelle, J. E. (2023). A systematic review of neuroimaging approaches to mapping language in individuals. *Journal of Neurolinguistics*, *68*, 101163. https://doi.org/10.1016/j.jneuroling.2023.101163

Bardovi-Harlig, K., & Dörnyei, Z. (1998). Do language learners recognize pragmatic violations? Pragmatic versus grammatical awareness in instructed L2 learning. *TESOL Quarterly*, *32*(2), 233-259. https://doi.org/10.2307/3587583

Barquero, L. A., Davis, N., & Cutting, L. E. (2014). Neuroimaging of reading intervention: a systematic review and activation likelihood estimate meta-analysis. *PloS one*, *9*(1), e83668. https://doi.org/10.1371/journal.pone.0083668

Bandettini, P. A. (2009). What's new in neuroimaging methods? *Annals of the New York Academy of Sciences*, *1156*(1), 260-293. https://doi.org/10.1111/j.1749-6632.2009.04420.x

*Barbeau, E. B., Chai, X. J., Chen, J. K., Soles, J., Berken, J., Baum, S., ... & Klein, D. (2017). The role of the left inferior parietal lobule in second language learning: An intensive language training fMRI study. *Neuropsychologia*, *98*, 169-176. https://doi.org/10.1016/j.neuropsychologia.2016.10.003

Bialystok, E. (2024). Bilingualism modifies cognition through adaptation, not transfer. *Trends in Cognitive Sciences*.*28*(11), 987-997. https://doi.org/10.1016/j.tics.2024.07.012

Blanken, T. F., Bathelt, J., Deserno, M. K., Voge, L., Borsboom, D., & Douw, L. (2021). Connecting brain and behavior in clinical neuroscience: A network approach. *Neuroscience & Biobehavioral Reviews*, *130*, 81-90. https://doi.org/10.1016/j.neubiorev.2021.07.027

Bolkan, S., & Goodboy, A. K. (2020). Instruction, example order, and student learning: Reducing extraneous cognitive load by providing structure for elaborated examples. *Communication Education*, *69*(3), 300-316. https://doi.org/10.1080/03634523.2019.1701196

*Buchweitz, A., Mason, R. A., Hasegawa, M., & Just, M. A. (2009). Japanese and English sentence reading comprehension and writing systems: An fMRI study of first and second language effects on brain activation. *Bilingualism: Language and Cognition*, *12*(2), 141-151. https://doi.org/10.1017/S1366728908003970

Butler, A., Hall, H., & Copnell, B. (2016). A guide to writing a qualitative systematic review protocol to enhance evidence-based practice in nursing and health care. *Worldviews on Evidence-Based Nursing*, *13*(3), 241-249. https://doi.org/10.1111/wvn.12134

Cattinelli, I., Borghese, N. A., Gallucci, M., & Paulesu, E. (2013). Reading the reading brain: a new meta-analysis of functional imaging data on reading. *Journal of neurolinguistics*, *26*(1), 214-238. https://doi.org/10.1016/j.jneuroling.2012.08.001

*Choi, S., Jang, K. E., Lee, Y., Song, H., Cha, H., Lee, H. J., ... & Chang, Y. (2018). Neural processing of lower-and upper-case text in second language learners of English: an fMRI study. *Language, Cognition and Neuroscience*, *33*(2), 165-174. https://doi.org/10.1080/23273798.2017.1384028

Chong, s. W., & Plonsky, L. (2024). A typology of secondary research in Applied Linguistics. *Applied Linguistics Review*, *15*(4), 1569-1594. https://orcid.org/0000-0002-4519-0544

Comstock, L. (2024). The role of research design in the reproducibility of L1 and L2 language networks: A review of bilingual neuroimaging meta-analyses. *Brain and Language*, *249*, 105377. https://doi.org/10.1016/j.bandl.2023.105377

Corrigan, N. M., Yarnykh, V. L., Huber, E., Zhao, T. C., & Kuhl, P. K. (2022). Brain myelination at 7 months of age predicts later language development. *Neuroimage*, *263*, 119641. https://doi.org/10.1016/j.neuroimage.2022.119641

Covey, L., Fiorentino, R., & Gabriele, A. (2024). Island sensitivity in L2 learners: Evidence from acceptability judgments and event-related potentials. *Second Language Research*, *40*(1), 19-50. https://doi.org/10.1177/02676583221116039

Dallas, A., DeDe, G., & Nicol, J. (2013). An event-related potential (ERP) investigation of filler-gap processing in native and second language speakers. *Language Learning*, *63*(4), 766–799. https://doi.org/10.1111/lang.12026

Davis, D. (2016). A practical overview of how to conduct a systematic review. *Nursing Standard*, *31*(12). https://doi.org/10.7748/ns.2016.e10316

Deldar, Z., Gevers-Montoro, C., Khatibi, A., & Ghazi-Saidi, L. (2020). The interaction between language and working memory: a systematic review of fMRI studies in the past two decades. *AIMS neuroscience*, *8*(1), 1. https://doi.org/10.3934/Neuroscience.2021001

Démonet, J. F., Thierry, G., & Cardebat, D. (2005). Renewal of the neurophysiology of language: functional neuroimaging. Physiological reviews, 85(1), 49 95. https://doi.org/10.1152/physrev.00049.2003

Desroches, A. S., Cone, N. E., Bolger, D. J., Bitan, T., Burman, D. D., & Booth, J. R. (2010). Children with reading difficulties show differences in brain regions associated with orthographic processing during spoken language processing. *Brain research*, *1356*, 73-84. https://doi.org/10.1016/j.brainres.2010.07.097

Du, B., Yang, Z., Wang, C., Li, Y., & Tao, S. (2023). Short-term training helps second-language learners read like native readers: An ERP study. *Brain and Language*, *239*, 105251. https://doi.org/10.1016/j.bandl.2023.105251

*Elgort, I., Perfetti, C. A., Rickles, B., & Stafura, J. Z. (2015). Contextual learning of L2 word meanings: Second language proficiency modulates behavioural and event-related brain potential (ERP) indicators of learning. *Language, Cognition and Neuroscience*, *30*(5), 506-528. https://doi.org/10.1080/23273798.2014.942673

Fedorenko, E., & Kanwisher, N. (2009). Neuroimaging of language: why hasn't a clearer picture emerged?. Language and linguistics compass, 3(4), 839-865. https://doi.org/10.1111/j.1749-818X.2009.00143.x

Friederici, A. D. (2011). The brain basis of language processing: from structure to function. *Physiological Reviews*, *91*(4), 1357-1392. https://doi.org/10.1152/physrev.00006.2011

Filipović, L. (2022). First language versus second language effect on memory for motion events: The role of language type and proficiency. *International Journal of Bilingualism*, *26*(1), 65-81. https://doi.org/10.1177/13670069211022863

Friederici, A. D., & Wartenburger, I. (2010). Language and brain. *Wiley Interdisciplinary Reviews: Cognitive Science*, *1*(2), 150-159. https://doi.org/10.1002/wcs.9

Fu, Y., Bermúdez-Margaretto, B., Beltrán, D., Huili, W., & Dominguez, A. (2024). Language proficiency modulates L2 orthographic learning mechanism: Evidence from event-related brain potentials in overt naming. *Studies in Second Language Acquisition*, 1-22. https://doi.org/10.1017/S0272263123000426

Gernsbacher, M. A., & Kaschak, M. P. (2003). Neuroimaging studies of language production and comprehension. *Annual review of psychology*, *54*(1), 91-114. https://doi.org/10.1146/annurev.psych.54.101601.145128

Godfroid, A., & Hopp, H. (Eds.). (2023). *The Routledge handbook of second language acquisition and psycholinguistics*. Routledge, Taylor & Francis Group.

*Grant, A. M., Fang, S. Y., & Li, P. (2015). Second language lexical development and cognitive control: A longitudinal fMRI study. *Brain and language*, *144*, 35-47. https://doi.org/10.1016/j.bandl.2015.03.010

Grey, S. (2023). Variability in native and nonnative language: An ERP study of semantic and grammar processing. *Studies in Second Language Acquisition*, *45*(1), 137-166. https://doi.org/10.1017/S0272263122000055

Guiral, J. A. (2024). Neuropsychological dimensions related to alterations of verbal self-monitoring neural networks in schizophrenic language: systematic review. *Frontiers in Psychiatry*, *15*, 1356726. https://doi.org/10.3389/fpsyt.2024.1356726

Herholz, K., Perani, D., Fazekas, F., Markus, H., Baumgartner, R. W., & Baron, J. C. (2001). A survey of neuroimaging research in European neurological departments. *European Journal of Neurology*, *8*(2), 111-117. https://doi.org/10.1046/j.1468-1331.2001.00206.x

Herringshaw, A. J., Ammons, C. J., DeRamus, T. P., & Kana, R. K. (2016). Hemispheric differences in language processing in autism spectrum disorders: A meta-analysis of neuroimaging studies. *Autism Research*, *9*(10), 1046-1057. https://doi.org/10.1002/aur.1599

Hu, W., & Luo, Y. (2024). Chinese English language learners' vocabulary retention: Investigating the effectiveness of neuro/metacognitive and socio-cultural strategies. *BMC psychology*, *12*(1), 113. https://doi.org/10.1186/s40359-024-01612-0

Kelsen, B. A., Sumich, A., Kasabov, N., Liang, S. H., & Wang, G. Y. (2022). What has social neuroscience learned from hyperscanning studies of spoken communication? A systematic review. *Neuroscience & Biobehavioral Reviews*, *132*, 1249-1262. https://doi.org/10.1016/j.neubiorev.2020.09.008

*Koyama, M. S., Stein, J. F., Stoodley, C. J., & Hansen, P. C. (2014). A cross-linguistic evaluation of script-specific effects on fMRI lateralization in late second language readers. *Frontiers in human neuroscience*, *8*, 249. https://doi.org/10.3389/fnhum.2014.00249

Kram, L., Ohlerth, A. K., Ille, S., Meyer, B., & Krieg, S. M. (2024). CompreTAP: Feasibility and reliability of a new language comprehension mapping task via preoperative navigated transcranial magnetic stimulation. *Cortex*, *171*, 347-369. https://doi.org/10.1016/j.cortex.2023.09.023

Kraus, B., Zinbarg, R., Braga, R. M., Nusslock, R., Mittal, V. A., & Gratton, C. (2023). Insights from personalized models of brain and behavior for identifying biomarkers in psychiatry. *Neuroscience & Biobehavioral Reviews*, *152*, 105259. https://doi.org/10.1016/j.neubiorev.2023.105259

Kumar, U., Dhanik, K., Mishra, M., Pandey, H. R., & Keshri, A. (2024). Mapping the unique neural engagement in deaf individuals during picture, word, and sign language processing: fMRI study. *Brain Imaging and Behavior*, 1-17. https://doi.org/10.1007/s11682-024-00878-7

Legault, J., Fang, S. Y., Lan, Y. J., & Li, P. (2019). Structural brain changes as a function of second language vocabulary training: Effects of learning context. *Brain and Cognition*, *134*, 90-102. https://doi.org/10.1016/j.bandc.2018.09.004

Li, Y. (2024). Can late language learners acquire second language grammar? Evidence from linguistic to neuroscience methods. *Frontiers in Psychology*, *15*, 1421072. https://doi.org/10.3389/fpsyg.2024.1421072

Li, P., Legault, J., & Litcofsky, K. A. (2014). Neuroplasticity as a function of second language learning: Anatomical changes in the human brain. *Cortex*, *58*, 301-324. https://doi.org/10.1016/j.cortex.2014.05.001

Lindell, A. K. (2006). In your right mind: Right hemisphere contributions to language processing and production. *Neuropsychology review*, *16*(3), 131–148. https://doi.org/10.1007/s11065-006-9011-9

Littlewood, W. (2004). Second language learning. In A. Davies, & C. Elder (Eds.), *The handbook of applied linguistics* (pp. 501–524). Malden, MA: Blackwell. https://doi.org/10.1002/9780470757000.ch20

*Liu, C., de Bruin, A., Jiao, L., Li, Z., & Wang, R. (2021a). Second language learning tunes the language control network: a longitudinal fMRI study. *Language, Cognition and Neuroscience*, *36*(4), 462-473. https://doi.org/10.1080/23273798.2020.1856898

*Liu, C., Jiao, L., Li, Z., Timmer, K., & Wang, R. (2021b). Language control network adapts to second language learning: A longitudinal rs-fMRI study. *Neuropsychologia*, *150*, 107688. https://doi.org/10.1016/j.neuropsychologia.2020.107688

Loosen, A. M., Kato, A., & Gu, X. (2024). Revisiting the role of computational neuroimaging in the era of integrative neuroscience. *Neuropsychopharmacology*, *50*(1), 103-113. https://doi.org/10.1038/s41386-024-01946-8

Lum, J. A., Clark, G. M., Bigelow, F. J., & Enticott, P. G. (2022). Resting state electroencephalography (EEG) correlates with children's language skills: Evidence from sentence repetition. *Brain and Language*, *230*, 105137. https://doi.org/10.1016/j.bandl.2022.105137

Makita, K., Yamazaki, M., Tanabe, H. C., Koike, T., Kochiyama, T., Yokokawa, H., ... & Sadato, N. (2013). A Functional Magnetic Resonance Imaging Study of Foreign-Language Vocabulary Learning Enhanced by Phonological Rehearsal: The Role of the Right Cerebellum and Left Fusiform Gyrus. *Mind, Brain, and Education*, *7*(4), 213-224. https://doi.org/10.1111/mbe.12029

Mancin, S., Sguanci, M., Anastasi, G., Godino, L., Cascio, A. L., Morenghi, E., ... & De Marinis, M. G. (2024). RETRACTED: A methodological framework for rigorous systematic reviews: Tailoring comprehensive analyses to clinicians and healthcare professionals. *Methods*, *225*, 38-43.

Michon, K. J., Khammash, D., Simmonite, M., Hamlin, A. M., & Polk, T. A. (2022). Person-specific and precision neuroimaging: Current methods and future directions. *NeuroImage*, *263*, 119589. https://doi.org/10.1016/j.neuroimage.2022.119589

* Midgley, K. J., Holcomb, P. J., & Grainger, J. (2009). Masked repetition and translation priming in second language learners: A window on the time-course of form and meaning activation using ERPs. *Psychophysiology*, *46*(3), 551-565. https://doi.org/10.1111/j.1469-8986.2009.00784.x

*Midgley, K. J., Holcomb, P. J., & Grainger, J. (2011). Effects of cognate status on word comprehension in second language learners: An ERP investigation. *Journal of cognitive neuroscience*, *23*(7), 1634-1647. https://doi.org/10.1162%2Fjocn.2010.21463

Momenian, M., Vaghefi, M., Sadeghi, H., Momtazi, S., & Meyer, L. (2024). Language prediction in monolingual and bilingual speakers: an EEG study. *Scientific Reports*, *14*(1), 6818. https://doi.org/10.1038/s41598-024-57426-y

Monroy-Sosa, A., Chakravarthi, S. S., Cortes-Contreras, A. P., Hernandez-Varela, M., Andres-Arrieta, V., Epping, A., & Rovin, R. A. (2021). The evolution of cerebral language localization: historical analysis and current trends. *World Neurosurgery*, *145*, 89-97. https://doi.org/10.1016/j.wneu.2020.09.028

Morgan-Short, K., Deng, Z., Brill-Schuetz, K. A., Faretta-Stutenberg, M., Wong, P. C., & Wong, F. C. (2015). A view of the neural representation of second language syntax through artificial language learning under implicit contexts of exposure. *Studies in Second Language Acquisition*, *37*(2), 383-419. https://doi.org/10.1017/S0272263115000030

Mueller, J. L. (2009). The influence of lexical familiarity on ERP responses during sentence comprehension in language learners. *Second Language Research*, *25*(1), 43-76. https://doi.org/10.1177/0267658308098996

Mulder, E., Van de Ven, M., Segers, E., & Verhoeven, L. (2019). Context, word, and student predictors in second language vocabulary learning. Applied Psycholinguistics, 40(1), 137-166. https://doi.org/10.1017/S0142716418000504

Mönte, T. F., Altenmüller, E., & Jäncke, L. (2002). The musician's brain as a model of neuroplasticity. *Nature Reviews Neuroscience, 3*, 473–478. https://doi.org/10.1038/nrn843

Nakagawa, E., Koike, T., Sumiya, M., Shimada, K., Makita, K., Yoshida, H., ... & Sadato, N. (2022). The Neural Correlates of Semantic and Grammatical Encoding During Sentence Production in a Second Language: Evidence From an fMRI Study Using Structural Priming. *Frontiers in Human Neuroscience*, *15*, 753245. https://doi.org/10.3389/fnhum.2021.753245

Osterhout, L., Poliakov, A., Inoue, K., McLaughlin, J., Valentine, G., Pitkanen, I., ... & Hirschensohn, J. (2008). Second-language learning and changes in the brain. *Journal of neurolinguistics*, *21*(6), 509-521. https://doi.org/10.1016/j.jneuroling.2008.01.001

Pinel, P., & Dehaene, S. (2010). Beyond hemispheric dominance: brain regions underlying the joint lateralization of language and arithmetic to the left hemisphere. *Journal of Cognitive Neuroscience*, *22*(1), 48-66. https://doi.org/10.1162/jocn.2009.21184

Pu, P., Chang, D. Y. S., & Wang, S. (2024). Incidental learning of collocations through different multimodal input: The role of learners' initial L2 proficiency. *System*, *125*, 103416. https://doi.org/10.1016/j.system.2024.103416

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., ... & Moher, D. (2021). The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ, 372*. https://doi.org/10.1136/bmj.n71

Parviainen, T., & Kujala, J. (2023). Event-Related Potentials (ERPs) and Event-Related Fields (ERFs). In *Language Electrified: Principles, Methods, and Future Perspectives of Investigation* (pp. 195-239). New York, NY: Springer US.

Perkins, K., & Zhang, L. J. (2024). The Effect of First Language Transfer on Second Language Acquisition and Learning: From Contrastive Analysis to Contemporary Neuroimaging. *RELC Journal*, *55*(1), 162-178. https://doi.org/10.1177/00336882221081894

Perpiñán, S. (2015). L2 grammar and L2 processing in the acquisition of Spanish prepositional relative clauses. *Bilingualism: Language and Cognition*, *18*(4), 577-596. https://doi.org/10.1017/S1366728914000583

Perret, M., Neige, C., Brunelin, J., & Mondino, M. (2024). Unraveling the brain mechanisms of source monitoring with non-invasive brain stimulation: A systematic review. *International Journal of Clinical and Health Psychology*, *24*(2), 100449. https://doi.org/10.1016/j.ijchp.2024.100449

Piller, I., & Bodis, A. (2024). Marking and unmarking the (non) native speaker through English language proficiency requirements for university admission. *Language in Society*, *53*(1), 1-23. https://doi.org/10.1017/S0047404522000689

Provost, S., Fourdain, S., Vannasing, P., Tremblay, J., Roger, K., Caron-Desrochers, L., ... & Gallagher, A. (2024). Language brain responses and neurodevelopmental outcome in preschoolers with congenital heart disease: A fNIRS study. *Neuropsychologia*, 108843. https://doi.org/10.1016/j.neuropsychologia.2024.108843

Pu, H., Holcomb, P. J., & Midgley, K. J. (2016). Neural changes underlying early stages of L2 vocabulary acquisition. *Journal of Neurolinguistics*, *40*, 55-65. https://doi.org/10.1016/j.jneuroling.2014.09.004

Quin-Conroy, J. E., Bayliss, D. M., Daniell, S. G., & Badcock, N. A. (2024). Patterns of language and visuospatial functional lateralization and cognitive ability: a systematic review. *Laterality*, *29*(1), 63-96. https://doi.org/10.1080/1357650X.2023.2263199

Raichle, M. E. (2009). A brief history of human brain mapping. *Trends in neurosciences*, *32*(2), 118-126. https://doi.org/10.1016/j.tins.2008.11.001

Rajimehr, R., Firoozi, A., Rafipoor, H., Abbasi, N., & Duncan, J. (2022). Complementary hemispheric lateralization of language and social processing in the human brain. *Cell Reports*, *41*(6). https://doi.org/10.1016/j.celrep.2022.111617

Rastelli, S. (2018). Neurolinguistics and second language teaching: A view from the crossroads. *Second Language Research*, *34*(1), 103-123. https://doi.org/10.1177/0267658316681377

Reber, P. J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia*, *51*(10), 2026-2042. https://doi.org/10.1016/j.neuropsychologia.2013.06.019

*Reiterer, S., Pereda, E., & Bhattacharya, J. (2009). Measuring second language proficiency with EEG synchronization: how functional cortical networks and hemispheric involvement differ as a function of proficiency level in second language speakers. *Second Language Research, 25*(1), 77-106. https://doi.org/10.1177/0267658308098997

Roberts, L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, *35*(2), 213-235. https://doi.org/10.1017/S0272263112000861

Rosselli, M., Ardila, A., Matute, E., & Vélez-Uribe, I. (2014). Language development across the life span: A neuropsychological/neuroimaging perspective. *Neuroscience Journal*, *2014*(1), 585237. https://doi.org/10.1155/2014/585237

Russo, C., & Senese, V. P. (2023). Functional near-infrared spectroscopy is a useful tool for multi-perspective psychobiological study of neurophysiological correlates of parenting behaviour. *European Journal of Neuroscience*, *57*(2), 258-284. https://doi.org/10.1111/ejn.15890

Sabourin, L. (2009). Neuroimaging and research into second language acquisition. *Second Language Research, 25*(1), 5-11. https://doi.org/10.1177/0267658308098994

Saldana, J. (2015). *The coding manual for qualitative researchers* (3rd ed). Sage Publications.

Savoy, R. L. (2001). History and future directions of human brain mapping and functional neuroimaging. *Acta Psychologica*, *107*(1-3), 9-42. https://doi.org/10.1016/s0001-6918(01)00018-x

Schlaggar, B. L., & Church, J. A. (2009). Functional neuroimaging insights into the development of skilled reading. *Current Directions in Psychological Science*, *18*(1), 21-26. https://doi.org/10.1111/j.1467-8721.2009.01599.x

Schiavo, J. H. (2019). PROSPERO: an international register of systematic review protocols. *Medical reference services quarterly*, *38*(2), 171-180. https://doi.org/10.1080/02763869.2019.1588072

Shaheen, N., Shaheen, A., Ramadan, A., Hefnawy, M. T., Ramadan, A., Ibrahim, I. A., ... & Flouty, O. (2023). Appraising systematic reviews: a comprehensive guide to ensuring validity and reliability. *Frontiers in research metrics and analytics*, *8*, 1268045. https://doi.org/10.3389/frma.2023.1268045

Sharma, R., & Meena, H. K. (2024). Emerging Trends in EEG Signal Processing: A Systematic Review. *SN Computer Science*, *5*(4), 1-14. https://doi.org/10.1007/s42979-024-02773-w

Sullivan, M. D., Janus, M., Moreno, S., Astheimer, L., & Bialystok, E. (2014). Early stage second-language learning improves executive control: Evidence from ERP. *Brain and language*, *139*, 84-98. https://doi.org/10.1016/j.bandl.2014.10.004

Sulpizio, S., Del Maschio, N., Fedeli, D., & Abutalebi, J. (2020). Bilingual language processing: A meta-analysis of functional neuroimaging studies. *Neuroscience & Biobehavioral Reviews*, *108*, 834-853. https://doi.org/10.1016/j.neubiorev.2019.12.014

Svaldi, C., Paquier, P., Keulen, S., van Elp, H., Catsman-Berrevoets, C., Kingma, A., ... & de Aguiar, V. (2024). Characterising the long-term language impairments of children following cerebellar tumour surgery by extracting psycholinguistic properties from spontaneous language. *The Cerebellum*, *23*(2), 523-544. https://doi.org/10.1007/s12311-023-01563-z

Tao, L., Wang, G., Zhu, M., & Cai, Q. (2021). Bilingualism and domain-general cognitive functions from a neural perspective: A systematic review. *Neuroscience & Biobehavioral Reviews*, *125*, 264-295. https://doi.org/10.1016/j.neubiorev.2021.02.029

Van Atteveldt, N., Van Kesteren, M. T., Braams, B., & Krabbendam, L. (2018). Neuroimaging of learning and development: improving ecological validity. *Frontline Learning Research*, *6*(3), 186. https://doi.org/10.14786/flr.v6i3.366

Von Rhein, D., Mennes, M., van Ewijk, H., Groenman, A. P., Zwiers, M. P., Oosterlaan, J., ... & Buitelaar, J. (2015). The NeuroIMAGE study: a prospective phenotypic, cognitive, genetic and MRI study in children with attention-deficit/hyperactivity disorder. Design and descriptives. *European child & adolescent psychiatry*, *24*, 265-281. https://doi.org/10.1007/s00787-014-0573-4

Wang, R., Ke, S., Zhang, Q., Zhou, K., Li, P., & Yang, J. (2020). Functional and structural neuroplasticity associated with second language proficiency: An MRI study of Chinese-English bilinguals. *Journal of Neurolinguistics*, *56*, 100940. https://doi.org/10.1016/j.jneuroling.2020.100940

*Wang, Y., Xue, G., Chen, C., Xue, F., & Dong, Q. (2007). Neural bases of asymmetric language switching in second-language learners: An ER-fMRI study. *NeuroImage*, *35*(2), 862-870. https://doi.org/10.1016/j.neuroimage.2006.09.054

Ware, C., Dautricourt, S., Gonneaud, J., & Chételat, G. (2021). Does second language learning promote neuroplasticity in aging? A systematic review of cognitive and neuroimaging studies. *Frontiers in Aging Neuroscience*, *13*, 706672. https://doi.org/10.3389/fnagi.2021.706672

Whelan, R. R. (2007). Neuroimaging of cognitive load in instructional multimedia. *Educational Research Review*, *2*(1), 1-12. https://doi.org/10.1016/j.edurev.2006.11.001

Wischmann, S., Kamper, N. R., Jantzen, L., Hammer, L., Reipur, D. B., Serafin, S., & Percy-Smith, L. (2024). Explaining neurological factors of hearing loss through digital technologies. *International Journal of Pediatric Otorhinolaryngology*, *176*, 111825. https://doi.org/10.1016/j.ijporl.2023.111825

*Xu, X., Pan, M., Dai, H., Zhang, H., & Lu, Y. (2019). How referential uncertainty is modulated by conjunctions: ERP evidence from advanced Chinese–English L2 learners and English L1 speakers. *Second Language Research*, *35*(2), 195-224. https://doi.org/10.1177/0267658318756948

*Yang, J., Gates, K. M., Molenaar, P., & Li, P. (2015). Neural changes underlying successful second language word learning: An fMRI study. *Journal of Neurolinguistics*, *33*, 29-49. http://dx.doi.org/10.1016/j.jneuroling.2014.09.004

Yongqi Gu, P. (2016). Questionnaires in language teaching research. *Language Teaching Research, 20*(5), 567-570. https://doi.org/10.1177/1362168816664001

Zhang, X., Yang, J., Wang, R., & Li, P. (2020). A neuroimaging study of semantic representation in first and second languages. *Language, Cognition and Neuroscience*, *35*(10), 1223-1238. https://doi.org/10.1080/23273798.2020.1738509

Zhang, Y., Zhao, J., Huang, H., Zhang, Z., Wu, S., Qiu, J., & Wu, Y. J. (2024). Neuroimaging evidence dissociates forced and free language selection during bilingual speech production. *Bilingualism: Language and Cognition*, 1-10. https://doi.org/10.1017/S1366728924000324

Zheng, Y., & Zhang, B. (2024). 25-year neuroimaging research on spoken language processing: a bibliometric analysis. *Frontiers in Human Neuroscience*, *18*, 1461505. https://doi.org/10.3389/fnhum.2024.1461505

8. Appendices

8.1. Appendix A

Characteristics of included studies

Author (year)	Country	Participants	Number of males	Number of females	Age (mean)	Main goals of the study	Evaluation Tools	Instruments
Reiterer et al., (2009)	DenmarkAustria and UK	University Student (English students and other fields)	19	0	24	Proficiency in speaking	EEG, Questionaries	spoken auditory speech input, short interviews, Comprehension questionnaire
Barbeau et al. (2017)	Canada and UK	Adults' French learners	NM (Total=14)	NM (Total=14)	NM	Neural changes during language learning	FMRI, Questionaries	Self-reported language competence, Language Expertise and Proficiency Questionnaire (LEAP-Q).
Grant et al., (2015)	USA	Late L2 Spanish learners	NM (Total=24)	NM (Total=24)	NM	Investigating individual differences in cognitive control and lexical architecture among late-L2 learners.	FMRI, Individual differences measure, working memory measure	Lexical decision task, semantic judgement task, Spanish proficiency test
Yang et al., (2015)	China and USA	American University student's learning Chinese as L2	19	20	NM	An fMRI scan was used to analyze the neural mechanisms contributing to tonal language learning in adult L2 learners prior to and following instruction.	FMRI	Stimuli and training tasks, vocabulary instruction
Sullivan et al., (2014)	Canada and USA	University students Spanish L2 learners	NM (Total=54)	NM (Total=54)	24.5	The study aimed to determine the emergence of neural processing modifications relevant to L2 learning during developmental stages.	ERP, questionary, test	Verbal fluency tasks, social background questionnaire, Peabody Picture Vocabulary Test, Kaufman Brief Intelligence Test
Midgley et al., (2011)	France	French L2 learners	9	33	20	An investigation was conducted to determine how cognate position affects word perception in L2 learners.	ERP	Cognate and non-cognate words.
Mueller, (2009)	Germany	Japanese L2 Learners	10	11	23.9	The purpose of this study is to examine ERP elements in adult language learners with respect to syntactic and thematic procedures.	ERP	Mini-Nihongo grammar game.
Liu,, et al., (2021)	China and Poland	University students	2	18	18.52	Language control processing and cognitive control	FMRI, Questionaries Oxford placement test	Language switching tasks
Liu et al., (2021)	China and UK	Pre-school students	2	20	18.25	Children's cognitive flexibility and creativity	FMRI, Questionaries Oxford placement test	Color shape switching tasks, Self-rated questionaries
Choi et al., (2018)	Korea	English learners	9	14	22.43	Examine the changes in brain activity and connectivity before and after exposure to VR	FMRI, English for International Communication (TOEIC), Nelson-Denny Reading Test (NDRT)	Super Lab programme, 40 trials of two conditions: lower-case trials and upper-case trials,
Buchweitz, et al. (2009)	USA and Japan	Native Japanese (L1) readers and English as a second language (L2) reader	NM (Total=9)	NM (Total=9)	27.4	The brain responds differently to different writing systems and reading comprehension in a second language	FMRI, Language background questionnaire	Edinburgh Inventory, Comprehension task, Two-clause target sentences

Wang et al., (2007)	USA and China	Later English learners	6	6	19.5	Examine neurocognitive models of lexical selection	ERP	Experimental tasks, self-rated language proficiency
Nakagawa et al., (2022)	Japan	English learners	14	16	22.7	Consideration should be given to the interaction between conceptualization and articulation as well as the automated nature of grammatical encoding.	FMRI	Cartoons, sentence fragment, stimuli from the actual experiment
Koyama et al., (2014)	UK and USA	Late L2 English learners	13	32	29.3	Studying the degree to which bilingual individuals, particularly late L2 learners, display functional lateralization in first and second language processing. The purpose of this research is to examine brain activation patterns during reading tasks in L1 and L2 languages.	FMRI, Wide Range Achievement Test III, Annett Handedness Questionnaire	Block design with alternating task, visual one-back matching task, Raven's Advanced Progressive Matrices
Du et al., (2023)	China	Late L2 English learners	24	15	24.3	This study examined whether learning skills specific to L2 reading could help learners overcome the limitations imposed by their native language.	ERP, The grammar subtest of the Oxford Placement Test, EEG recording	English word-reading. A total of 200 words were selected from China's national curriculum for college English.
Dallas et al., (2013)	USA	Late L2 English learners	12	27	24.5	Investigate the real-time processing of filler-gap dependency sentences by late-learning speakers of English as a second language (L2).	Shipley Vocabulary Test, Wechsler Adult Intelligence Scale (WAIS), EEG recording	E-prime experimental software, the sentence processing task, presentation list contained 60 experimental sentences.
Midgley et al., (2009)	USA France	L2 English Learners	4	32	20.3	Analyze the time sequence in which form and meaning are perceived during word recognition, along with the role played by semantic representations. Researchers will be able to gain a deeper understanding of how these elements are activated as a result.	ERP masked repetition priming, EEG recording, EEG recording, Questionnaire	within- and cross-language primes, Visual stimuli
Elgort et al., (2015)	New Zealand USA	L2 English learners	NM (Total=26)	NM (Total=26)	NM	A study was conducted to examine the impact of accidental situational L2 word learning on lexical competence in L2.	Vocabulary Size Test, FMRI, ERP	L2 lexical proficiency, including key words in three sentences with high constraint levels.
Xu et al., (2019)	China	Both native (L1) readers and second language (L2) readers	NM (Total=49)	NM (Total=49)	NM	Analyze how native and non-native speakers interpret ambiguous pronouns with different conjunctions	ERP	60 two clause filler items, stimuli included protagonists while the other included only protagonists (equally for male and females)

8.2. Appendix B

Outcomes of included articles

Author (year)	Authors conclusion	Statistically significant	Outcomes			Limitations
			Positive	No difference	Negative	
Reiterer et al., (2009)	In the low proficiency group, the right hemisphere is more involved, especially in the L2, than in the superior performance group.	Yes		Based on other psychological indicators (self-reported attention, work load, sympathy for the speaker, and interest in the topic), no significant differences were found between the two groups.	A negative relationship is generally observed between proficiency level and synchronization density.	NM
Barbeau et al. (2017)	It was found that increased involvement of the left IPL was correlated with improved L2 function.	Yes	Reading, speaking, and comprehension skills in L2 were significantly improved in trainees.			Participants in this study were limited to a small number of adults.
Grant et al., (2015)	The neural processes involved in L2 lexical generation become increasingly similar to those involved in native language lexical generation with continued exposure.	Yes	An examination of the role of cognitive control in L2 learning and conceptual frameworks for lexical representation by late L2 learners.			It is impossible to continuously observe the evolution of lexical computation structures.
Yang et al., (2015)	It was demonstrated in the study that competent L2 participants exhibited decreased brain activity in specific brain areas following training, demonstrating improved language learning capacity.	Yes	Improved neural function for language use and argument clarification in proficient learners, indicating enhanced competence.			Inadequate focus on sound distinction competence prior to instruction, which may affect behavioral performance.
Sullivan et al., (2014)	After brief instruction in L2, significant changes were observed in neural activities related to contrasting tasks.	Yes	Spanish learners exhibited modifications in task execution after instruction.		Electrophysiological changes were not preceded by behavioral changes.	There is a shortage of bilingual students and a focus on L2 learning's initial phases.
Midgley et al., (2011)	The N400 magnitudes of cognate words in both L1 and L2 languages were lower compared with those of noncognates.	Yes	Increasing comprehension of how cognate identity affects word perception in L2 trainees.		Restricted generalization due to particular participant characteristics and language mixtures.	Limited to L1 English trainees of French, consequently restricting broader usability.
Mueller, (2009)	Word classification problems resulted in an initial failure accompanied by a P600 in commonly used sentences.	Yes	During training, trainees demonstrated a native-like understanding of well-known sentences.		Unknown words led to greater computational requirements and difficulties in data manipulation.	Differentiations among individuals may have different effects.
Liu, et al., (2021)	A long-term study of language in a classroom modulates the resting-state connectivity of the language control network, suggesting neural adaptations due to language learning.	Yes	The resting-state connectivity of the language control network is modulated by language learning, suggesting that neural adaptations are involved in this process.			In the process of language control, cognitive control plays an important role.
Liu et al., (2021)	In areas responsible for language control, long periods of classroom instruction in L2 can induce significant neuroplastic changes.	Yes	Learning L2 in the classroom significantly increased the connectivity between the dACC and LCN, and this increase was significantly correlated with behavioral language improvement switching cost		In the post-learning session, the language switch cost was significantly lower than in the pre-learning session.	A majority of the participants in the study were female, which may limit the generalizability of the results.
Choi et al., (2018)	According to the findings, uppercase text reduces reading comprehension	Yes	provide the		There was a negative effect of upper-case text on the reading	While the lower-case text set resulted in statistically significantly

	because it demands readers to use more lower-order neurocognitive resources.		first neurocognitive evidence with regard to the automaticity theory		comprehension of L2 learners, which is consistent with studies involving L1 learners.	higher response accuracy and shorter reaction times than the upper-case text set, the computed effect sizes were small.
Buchweitz et al., (2009)	Studies have demonstrated that cognitive responses can differ when processing at the sentence level within a given language, even if processing at the word level is used.	Yes	The reading comprehension of L2 was significantly more time-consuming than that of L1.			Based on the results, the L2 is systematically activated in relation to the L1 writing systems.
Wang et al., (2007)	Language switching exhibits different neural correlates depending on the direction of the switch, and no specific brain region appears to serve as a "language switch".	Yes	A comparison of the brain regions activated by forward switching compared to backward switching.			NM
Nakagawa et al., (2022)	As a result of the overload on the linguistic operation of formulating sentences and its associated executive function, it is difficult to form sentences in L2.	Yes	This study examined the process of speech production, particularly the stages of conceptualization, formulation, and articulation			Data from L1 English speakers were not included in the current study.
Koyama et al., (2014)	L2 readers are more likely to engage in visuo-spatial processing in the right hemisphere as a result of weaker leftward lateralization, rather than as a result of additional effort exhibited by the right hemisphere.	Yes	In word reading tests, there were varying accuracy scores between the L1 and L2 groups, indicating differences in cognitive performance			NM
Du et al., (2023)	If adults are trained in critical skills for reading a second language at appropriate dosages, they may be able to learn the skill of reading with native-like neural responses.	Yes	The findings of this study demonstrate that only short-term intensive and appropriate instruction specific to English letters and sound connections is effective in altering the neural responses of second-language learners.			Training benefits were not examined in this study for a period of time.
Dallas et al., 2013)	The L2 group did not process filler gap sentences in the same manner as the L1 group.	Yes	Speakers of second languages (L2) demonstrate a similar sensitivity to plausibility variations to speakers of first languages (L1).	A more native-like processing of filler-gap constructions was not associated with working memory.	Filler-gap sentences are not sensitive to plausibility variations among L2 speakers	NM
Midgley et al., (2009)	Form representations of non-cognate translation equivalents are intensively processed by L2 learners during on-going form-level processing of printed words, without any facilitative interaction between form equivalents.	Yes	There were significant interactions between repetition and the between-hemisphere language condition.	The interaction between repetition and priming type		NM
Elgort et al., (2015)	Newly learned L2 words can be recognized based both on episodic and lexical semantic knowledge. Depending on the learner's lexical proficiency in L2 and the context in which the word is used, it may vary.	Yes	In order to compensate for the lower processing efficiency of the pre-frontal cortex, bilinguals with lower proficiency engage the pre-frontal cortex when processing L2 words	Relatedness and proficiency did not interact		NM
Xu et al., (2019)	There is a greater likelihood of referential uncertainty for non-native speakers in sentences conjoined by conjunctions with complex semantics.	Yes	The effect of pronouns with conjunctions, such as "although," was greater for native speakers than for non-native speakers.		Referential uncertainty is more likely to occur when sentences are conjoined by conjunctions with complex semantics	NM

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 108-120

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.9117.1115

Interplay of Spiritual Intelligence and Foreign Language Enjoyment in Willingness to Communicate: A Case of High School EFL Students

Farzad Sabeki[1]
Ali Beikian[2*]
Elahe Alikarimi[3]
Rezvaneh Karimzadeh[4]

[1]Ph.D. Candidate in Applied Linguistics, Faculty of Management and Humanities, Department of English, Chabahar Maritime University, Chabahar, Iran

[*2]Assistant Professor of Translation Studies, Faculty of Management and Humanities, Department of English, Chabahar Maritime University, Chabahar, Iran

[3]M.A. in Applied Linguistics, Rafsanjan Office of General Education, Ministry of Education of Rafsanjan, Rafsanjan, Iran

[4]M.A. in Applied Linguistics, Chabahar Office of General Education, Ministry of Education of Chabahar, Chabahar, Iran

ARTICLE HISTORY

Received: 16 December 2024
Revised: 05 January 2025
Accepted: 10 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: a_beikian@cmu.ac.ir

ABSTRACT

Language learning is a multifaceted process influenced by a range of cognitive, emotional, and psychological factors. This study explores the relationships among spiritual intelligence (SI), foreign language enjoyment (FLE), and willingness to communicate (WTC) among Iranian high school learners of English as a Foreign Language (EFL). A quantitative approach was employed, utilizing structural equation modeling (SEM) to analyze data collected from 280 high school students in Kerman Province, Iran. The participants, selected through convenience sampling, included 78 males and 202 females, with ages ranging from 16 to 19 years. Data were gathered using three validated instruments, i.e., a WTC questionnaire, an SI scale, and an FLE scale. The findings reveal robust positive correlations among the three constructs, with both SI and FLE emerging as significant predictors of WTC. Additionally, a strong interdependence was observed between SI and FLE, highlighting their mutually reinforcing roles in shaping learners' communicative tendencies. These results underscore the importance of integrating psychological and emotional dimensions into language education, offering valuable insights for educators seeking to enhance learners' willingness to engage in communication.

KEYWORDS: Foreign language enjoyment; Iranian EFL learners; Spiritual intelligence; Structural equation modeling; Willingness to communicate

1. Introduction

The process of acquiring a new language is inherently complex, characterized by the intricate interplay of multifaceted elements that collectively shape learners' developmental trajectories throughout their language-learning journey (Benevene et al., 2020). This complexity underscores the necessity for a holistic understanding of the diverse components integral to language acquisition, which are deeply intertwined with the psychological and emotional dimensions of the learners themselves. Recent scholarly investigations have illuminated the pivotal role of these dimensions, particularly SI, FLE, and WTC, in determining the success of language acquisition (Derakhshan et al., 2022; Zhang et al., 2024; Zhao et al., 2024). Grounded in Gardner's (1983) theory of multiple

Cite this article: Sabeki, F. , Beikian, A. , Alikarimi, E. and Karimzadeh, R. (2025). Interplay of spiritual intelligence and foreign language enjoyment in willingness to communicate: A case of high school EFL students. *Applied Linguistics Inquiry*, *3*(1), 108-120. doi: 10.22077/ali.2025.9117.1115

intelligences, which posits that learners possess a spectrum of intelligence, research has consistently examined the domains of cognitive and emotional intelligence and their respective impacts on various facets of foreign language acquisition, including learner autonomy, academic achievement, and motivation (MacIntyre et al., 2019).

Building on Gardner's (1983) theoretical foundation, Zohar and Marshall (2001) introduced the concept of SI, marking a significant advancement in the understanding of intelligence beyond traditional cognitive frameworks. SI represents an autonomous domain that necessitates a novel perspective on the relationship between one's inner self, spirituality, and the external world (Vaughan, 2002). It encompasses the capacity to address complex problems, particularly those involving existential questions about the significance and value of diverse lifestyles (Zohar & Marshall, 2004). Unlike conventional intelligence paradigms, SI transcends cognitive abilities and normative psychological growth, bridging the personal and transpersonal realms and fostering a connection between the self and the spiritual dimension (Estaji & Pourmostafa, 2020). Despite its conceptual richness, however, the integration of SI into language learning research remains underexplored.

While the social and interaction-oriented dimensions of foreign language learning have garnered substantial attention, the critical role of positive emotions in this context has been comparatively overlooked (Dornyei & Ryan, 2015). The burgeoning field of positive psychology has catalyzed a paradigmatic shift toward examining the impact of positive emotions on language acquisition, emphasizing their transformative potential in enhancing learning outcomes (Wang & Wang, 2025). Within this framework, enjoyment emerges as a salient construct that not only fosters engagement in communicative activities but also amplifies learners' WTC, thereby contributing to overall linguistic proficiency (Botes et al., 2022). FLE, defined as a positive emotional state that promotes language learning through playful and exploratory tasks, has been identified as a key determinant of productive skill development (Akkas et al., 2022; Wang et al., 2021). Learners who effectively regulate their emotions to enhance FLE demonstrate heightened motivation to participate in communicative practices during language classes (Zarrinabadi & Pawlak, 2021).

The centrality of communication in second language acquisition (SLA) cannot be overstated, as underscored by recent studies highlighting its indispensable role in fostering genuine communicative competence (Cao, 2014). The advent of positive psychology in applied linguistics has further accentuated the affective turn, with emotions such as FLE playing a crucial role in shaping learners' WTC (Yu & Ma, 2024). Initially conceptualized as the readiness to initiate communication spontaneously in a second language at a specific moment with a particular interlocutor or group (MacIntyre et al., 1998), WTC has since emerged as a recurring theme in SLA research. Scholars have increasingly investigated the factors influencing WTC, emphasizing its importance as a predictor of successful language learning (Dewaele & Dewaele, 2018; Kruk, 2022; Nugroho, 2021; Yetkin & ÖZER, 2022). As MacIntyre and Vincze (2017) argue, the primary objective of foreign language education is to cultivate verbal self-expression, which serves as a catalyst for authentic communication and enhances language proficiency within educational settings.

Despite the growing body of interdisciplinary research on psychological factors in language learning, empirical investigations into the interplay between SI, FLE, and WTC among elementary and pre-intermediate Iranian EFL learners remain conspicuously sparse. To the best of the authors' knowledge, no prior study has simultaneously analyzed the roles of these constructs, underscoring the pressing need for further exploration of their combined influence. Addressing this lacuna, the present study seeks to contribute to the extant literature by employing a structural equation modeling (SEM) approach to examine the relationships between SI, FLE, and WTC among pre-intermediate Iranian EFL learners.

The significance of this study lies in its potential to transform the way educators and researchers approach the development of communicative competence in language learning. By empirically validating the interplay between SI, FLE, and WTC, the study provides a foundation for understanding how cognitive, emotional, and existential factors work together to shape learners' willingness to engage in communication. This holistic perspective is particularly vital in the Iranian EFL context, where traditional teaching methods have often overlooked the role of spiritual and emotional dimensions in fostering communicative behaviors. Furthermore, the findings are expected to offer actionable insights for educators, enabling them to design interventions that nurture learners' SI and enhance FLE.

2. Literature review

The review begins by tracing the theoretical underpinnings of SI, highlighting its multidimensional nature and its capacity to foster existential reflection and problem-solving. It then transitions to the role of FLE, a product of positive psychology, which underscores the importance of enjoyment and emotional engagement in language learning environments. Finally, the review examines WTC, a critical behavioral outcome influenced by both internal traits and external factors. By synthesizing prior studies on these constructs, this review sets the stage for understanding their interconnectedness and lays the groundwork for investigating their combined influence on communicative behaviors in second-language acquisition.

2.1. Spiritual intelligence (SI)

The role of individual differences in second-language acquisition has long been attributed to intelligence, a foundational concept in psychology that is universal among humans. Gardner (1983) defined intelligence as the capacity to solve problems within cultural contexts, expanding the traditional understanding of intelligence beyond intelligence quotient (IQ) and emotional quotient (EQ) through his theory of Multiple Intelligences. This theoretical framework laid the groundwork for the conceptualization of SI by Zohar in 1997, which sought to distinguish human cognition from other forms of life in the late 20th and early 21st centuries.

Despite its significance, SI remains challenging to define or operationalize within psychological research. Wigglesworth (2003) posited that spirituality represents an intrinsic human need that facilitates problem-solving and connects individuals to something transcendent beyond the physical realm. Emmons (2000) characterized SI as the ability to overcome challenges and achieve goals, emphasizing its internal connection to the mind, spirit, and external world. Zohar (2005) further elaborated on SI, describing it as the capacity to seek higher meanings, values, and core objectives while fostering imaginative lifestyles and exploring unconscious aspects of the self. Vaughan (2002) similarly linked SI to the ability to engage in deep existential reflection and navigate various levels of consciousness. King (2009) identified four key dimensions of SI: critical existential contemplation, the development of personal meaning, recognition of the transcendent, and the expansion of consciousness.

Advocates of SI, such as Hassan (2009) and Zohar (2012), argue that it fosters a learning environment conducive to maximizing learners' potential by focusing on significant life experiences, awareness, and analytical thinking. This necessitated the development of reliable tools to measure SI. For instance, Azadi et al. (2022) employed a sequential mixed-methods approach to construct and validate a questionnaire assessing SI among EFL learners. Their study surveyed 360 participants using a 27-item questionnaire developed through a literature review and interviews with 22 EFL learners. The findings demonstrated the efficacy of the new SI tool and revealed correlations between SI, gender, and educational level. Other studies have explored SI's relationship with language learning. Azizi and Zamaniyan (2013) found that high SI enables students to employ vocabulary learning strategies more effectively, enhancing their overall learning outcomes. Similarly, Aghaei et al. (2014) identified significant positive correlations between SI, language proficiency, and self-esteem among Iranian high school female students. Babazadeh et al. (2018) further examined the interplay between SI, multiple intelligences, and language learning strategies, revealing moderate positive relationships between these constructs.

2.2. Foreign language enjoyment

The introduction of positive psychology into language learning research by Mercer and MacIntyre (2014) marked a paradigm shift toward examining the role of positive emotions, particularly FLE, in language acquisition. Positive emotions are believed to enhance language learning by increasing awareness, improving classroom focus, fostering resilience, and strengthening personal resources (Fredrickson, 2003, 2006; MacIntyre & Gregersen, 2012). Language educators recognize the importance of cultivating motivation, determination, and resilience alongside positive emotions to sustain the language learning process (Mercer & MacIntyre, 2014). Fredrickson's (2001) broaden-and-build theory suggests that pleasure and playfulness in language learning promote social connections and cognitive growth.

Dewaele and MacIntyre (2014) conducted pioneering research on the relationship between FLE and foreign language classroom anxiety (FLCA), surveying 1,746 multilingual participants globally. Their findings revealed significantly higher levels of FLE compared to FLCA, indicating that the absence of enjoyment does not necessarily correlate with heightened anxiety. However, a strong inverse relationship was observed between FLE and FLCA, though statistical evidence did not support their distinctiveness. Dewaele et al. (2018) extended this line of inquiry by examining learner-internal and teacher/classroom-specific factors influencing FLE and FLCA among 189 British high school students. The results showed a small but significant negative correlation between FLE and FLCA, with learner-related factors (e.g., age, gender, language proficiency, attitudes) exerting a stronger influence than teacher-related variables.

Recent studies have further explored the implications of FLE in diverse educational contexts. Ma et al. (2024) investigated how enjoyment and burnout mediate the relationship between teacher-student interactions and foreign language achievement among 4,900 Chinese learners. Their findings highlighted gender differences in mediation effects, with boys exhibiting more pronounced mediation than girls. Pan and Zhang (2023) examined the development of FLE and FLA over 14 weeks among 55 college students, revealing that extraversion significantly influenced both constructs. Alrabai (2024) utilized PP to explore the interplay between emotions and motivation among 328 Saudi learners, identifying enjoyment as a mediator between motivation and L2 WTC (L2WTC). Wang and Wang (2025) analyzed the emotional experiences of 27 dual-language students, uncovering three primary attitudes: balanced enjoyment, activity-induced boredom, and low emotion regulation.

2.3. Willingness to communicate

The challenge of motivating learners to engage in communication is a persistent concern in language classrooms worldwide. The concept of WTC emerged as a personality trait linked to self-examination and self-esteem (Burgoon, 1976). McCroskey (1992) defined WTC as the likelihood of initiating communication when given the freedom to do so. While some researchers view WTC as a stable characteristic (McCroskey & Richmond, 1990), others argue that it is influenced by individual traits such as self-confidence, motivation, introversion/extroversion, perceived communication ability, communication apprehension, and self-esteem (Fallah, 2014; MacIntyre, 1994; Peng, 2007; Yu et al., 2011).

Tavakoli and Davoudi (2017) developed and validated a survey to assess oral WTC among 117 Iranian EFL students, identifying three dimensions: WTC with instructors, peers, and unfamiliar individuals. Their findings indicated that interlocutor type significantly influenced WTC, with students preferring communication with instructors or in front of the class rather than with peers in pairs or small groups. Sato and Dussuel Lam (2021) explored the impact of metacognitive instruction (MI) on WTC, revealing that MI enhanced target language use and equitable speaking turn distribution, though it did not significantly affect WTC itself.

Yashima et al. (2018) examined the interaction between learner characteristics and contextual factors in shaping state WTC among 21 Japanese EFL university students. Their findings underscored the influence of personality, skill level, peer responses, and

group communication patterns on WTC. Recent studies have also highlighted the role of emotions in WTC. Yu and Ma (2024) demonstrated that FLE, directly and indirectly, influences L2 WTC through mediators like Growth Language Mindset (GLM) and grit. Yang and Lin (2024) assessed the impact of FLE and classroom environment on WTC among Chinese EFL learners, revealing significant increases in WTC over a semester.

Despite the wealth of research on individual factors influencing language learning, a significant gap remains in the simultaneous examination of SI, FLE, and WTC within a unified framework. Existing studies have predominantly explored these constructs in isolation or limited combinations, failing to capture their dynamic interplay. For instance, while Babazadeh et al. (2018) examined the relationship between SI and language learning strategies, they overlooked its direct impact on communicative behaviors like WTC. Similarly, Dewaele and MacIntyre's (2014) pioneering work on FLE primarily focused on its inverse relationship with anxiety, neglecting its potential synergy with other cognitive and existential factors. Furthermore, Tavakoli and Davoudi's (2017) exploration of WTC dimensions did not account for the influence of deeper constructs such as SI or the mediating role of positive emotions like FLE. This fragmented approach has left unanswered questions about how these constructs collectively shape communicative readiness. Addressing this gap, the present study investigates the interplay between SI, FLE, and WTC among Iranian EFL learners, employing SEM to elucidate their complex relationships and contribute to a more holistic understanding of language learning outcomes.

Therefore, the present study seeks to answer the following research questions:

RQ1. Is there any significant relationship between spiritual intelligence, foreign language enjoyment, and willingness to communicate among Iranian high school EFL learners?

RQ2. Do spiritual intelligence and foreign language enjoyment significantly predict willingness to communicate among Iranian high school EFL learners?

3. Methodology

3.1. Research design

This study adopted a quantitative approach with a correlational design, leveraging numerical data and statistical analyses to address the research questions. The primary objective was to explore the relationships and predictive capacities of SI and FLE on WTC among Iranian high school EFL learners. By employing advanced statistical techniques, including Pearson correlation analysis and SEM, this study aimed to provide robust insights into the interplay between these constructs. The correlational design allows for the examination of relationships between variables without manipulating them, ensuring an ethical and practical approach to understanding their dynamic interactions in the context of language learning.

3.2. Participants

The sample comprised 280 high school students (females: n = 202, 72%; males: n = 78, 28%; Mean age = 16.88, SD = 0.79) enrolled in state schools across Kerman Province, Iran Convenience sampling was utilized to recruit participants, ensuring accessibility and feasibility within the educational context. All participants were native Persian speakers, ensuring linguistic homogeneity across the sample. Representing diverse socioeconomic backgrounds, the students were drawn from both urban and rural areas within Kerman Province, providing a broad and representative perspective of the population under study. Based on teacher reports, the students' English proficiency levels varied between elementary and pre-intermediate. This diverse demographic composition enhances the generalizability of the findings to similar educational contexts.

3.3. Instruments

3.3.1. Willingness to communicate questionnaire

WTC was assessed using WTC questionnaire, developed by Lee and Hsieh (2019), which is comprised of three sub-sections namely, WTC inside the class, WTC outside the class, and WTC in digital contexts. Each subsection has 4 items scored on a Likert scale ranging from 1 (definitely not willing) to 5 (definitely willing). The questionnaire provides a comprehensive measure of learners' communicative tendencies across different settings, capturing both face-to-face and digital interactions. To ensure the reliability of the instrument, Cronbach's alpha was calculated, yielding a coefficient of 0.86, which falls within the acceptable range for internal consistency. This suggests that the scale is both reliable and suitable for assessing WTC among the participants in this study.

3.3.2. Spiritual intelligence questionnaire

To determine spiritual intelligence among participants, spiritual intelligence scale developed and validated by Azadi et al. (2022) was used. This well-structured instrument comprises four sub-scales: (1) learning English for personal, social, and educational benefits, (2) learning English for personal, social, and academic achievement, (3) learning English to promote religious values, and (4) learning English for intercultural communication. The scale consists of 27 items, each measured on a five-point Likert

scale ranging from strongly agree to strongly disagree, allowing for nuanced responses that capture participants' perspectives on the role of English learning in relation to spiritual intelligence. A sample item from this scale is "I respect all people with different languages and beliefs". The reliability of the scale was assessed using Cronbach's alpha, which yielded a value of 0.95, indicating excellent internal consistency and demonstrating the instrument's robustness for measuring spiritual intelligence in diverse contexts. The high reliability, coupled with the scale's comprehensive sub-scales, ensures that it effectively captures the multifaceted nature of spiritual intelligence, particularly in the context of language learning and intercultural communication.

3.3.3. Foreign language enjoyment scale

The FLE scale, adopted from Dewaele and Macintyre's (2014) study, was employed to measure participants' levels of enjoyment in learning a foreign language. This well-established scale consists of 21 items, each rated on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree), allowing for a detailed assessment of various dimensions of foreign language enjoyment. The instrument has two main sub-scales including private and social. Items 1-13 measure FLE private and items 14-21 measure FLE social aspects. The scale captures a wide range of emotional and experiential aspects related to language learning, such as positive classroom experiences, personal satisfaction, and interactions with peers and teachers. A sample item from this questionnaire is "I've learnt interesting things". In the present study, the reliability of the FLE scale was evaluated using Cronbach's alpha, which yielded a value of 0.91, indicating excellent internal consistency and demonstrating the instrument's robustness for measuring foreign language enjoyment.

3.4. Data collection procedures

In the initial phase, the English questionnaires were translated into Persian to eliminate potential linguistic barriers and ensure clarity for participants. The translations were meticulously reviewed and double-checked by the researchers to ensure accuracy and cultural appropriateness. No significant issues were identified during this process. Subsequently, the questionnaires were digitized and uploaded to Google Forms to facilitate easy distribution and accessibility. A unique link to access the forms was generated and shared with participants via WhatsApp groups. To expand the reach and diversity of the participant pool, the researchers collaborated with colleagues in various cities, who distributed the link to their respective students. At the beginning of the online questionnaire, an informed consent section was included to emphasize that participation was entirely voluntary. Participants were required to acknowledge and agree to this section before proceeding. To ensure transparency and trust, teachers verbally assured students that the collected data would be used solely for research purposes and would remain strictly confidential. Ethical approval for the study was granted by the Ethics Committee of the Department of Education in Kerman Province, ensuring compliance with ethical standards throughout the research process. These procedures ensured both the integrity of the research process and the ethical treatment of participants.

3.5. Data analysis

The collected data were systematically coded and entered into SPSS software, version 27, for analysis. Following data entry, the dataset was carefully inspected for missing values and potential outliers to ensure data integrity and accuracy. The Kolmogorov-Smirnov test was conducted to assess the normality of the data, confirming that the data were normally distributed and thus suitable for parametric statistical tests. To address the first research question, Pearson correlation analysis was performed to examine the relationships between the study variables. Subsequently, SEM was conducted using AMOS 24 software to determine the impact of learners' SI and FLE on their WTC. SEM allowed for the simultaneous examination of direct and indirect effects, providing a nuanced understanding of the complex relationships among the constructs. This rigorous methodological framework ensures the reliability and validity of the findings, contributing to a deeper understanding of the interplay between SI, FLE, and WTC in the Iranian EFL context.

4. Results

4.1. Results of Kolmogorov-Smirnov test, descriptive statistics and correlations among variables

The present study sought to explore the intricate relationships between SI, FLE, and WTC among Iranian high school EFL learners. Specifically, the investigation aimed to address two pivotal inquiries: first, whether there exists a significant relationship between learners' SI, their experience of FLE, and their WTC; and second, whether SI and FLE serve as significant predictors of WTC. The findings presented in this section shed light on these questions, offering empirical insights into how these constructs interrelate and collectively influence communicative tendencies within the context of English language learning.

Before delving into the first research question, it is imperative to establish the normality of the data, a foundational prerequisite for robust statistical analysis. To this end, the Kolmogorov-Smirnov (K-S) test was rigorously employed to scrutinize whether the data distribution exhibited significant deviations from a normal distribution. As Field (2024) contends, a non-significant p-value—exceeding the conventional threshold of 0.05—is indicative of data that adheres to the assumptions of normality. The findings of the K-S test, as delineated in Table 1, serve as a critical juncture in validating the integrity of the subsequent analyses. Any deviation from normality would have undermined the reliability of the inferential statistics; thus, the confirmation of normality not only fortifies the methodological rigor of this study but also ensures the validity of the ensuing

interpretations.

Table 1. Tests of normality for different variables

	Kolmogorov-Smirnov		
	Statistics	df	Sig.
WTC	.041	280	.200
SI	.038	280	.200
FLE	.054	280	.045

Table 1 unveils a compelling distinction in the distribution patterns of the study variables. Both WTC and SI demonstrate adherence to a normal distribution, as evidenced by their non-significant p-values, thereby satisfying the statistical assumptions required for robust analysis. Conversely, the variable FLE presents a more contentious scenario, with a p-value of 0.045 that narrowly breaches the conventional threshold of 0.05. This marginal deviation from normality, while statistically detectable, remains sufficiently close to the cutoff to warrant cautious acceptance within the context of preliminary analysis. Such a finding underscores the importance of exercising analytical prudence; however, given the proximity of the p-value to the threshold, this slight deviation is unlikely to exert a substantial impact on the overall integrity of the results. Thus, the data collectively provide a defensible basis for proceeding with subsequent analyses, albeit with an acknowledgment of this nuanced limitation.

Table 2 presents a comprehensive overview of the descriptive statistics, offering critical insights into the dataset. The table includes the number of participants, means, and standard deviations for each variable under investigation, providing a clear snapshot of the central tendencies and variability within the data. Additionally, the correlations among the variables are delineated, enabling an initial assessment of the relationships between SI, FLE, and WTC. These statistical measures serve as foundational elements for understanding the patterns and interconnections that underpin the research questions, thereby setting the stage for more nuanced inferential analyses.

Table 2. Descriptive statistics and correlations for the variables

Variable	N	M	SD	Min	Max	1	2	3
1.WTC	280	39.91	9.96	15	60	-	.690**	.649**
2.SI	280	97.63	21.06	29	135		-	.797**
3.FLE	280	79.52	13.39	39	105			-

*p < .05. **p < .01

The data in Table 2 reveal compelling insights into the central tendencies of the variables under investigation. Notably, the mean score for WTC among EFL learners stands at 39.91, while SI and FLE exhibit mean scores of 97.63 and 79.52, respectively. These figures underscore the relative strengths of these constructs within the sample, suggesting that learners' experiences of SI and FLE are pronounced and potentially influential in shaping their communicative behaviors. However, such descriptive statistics alone cannot substantiate the hypothesized relationships between these variables; they merely set the stage for a more rigorous interrogation through inferential analysis.

In reply to the first research question, interrogating the existence of significant relationships between SI, FLE, and WTC, a Pearson correlation analysis was conducted. The findings, as outlined in Table 2, present an irrefutable case for the interconnectedness of these constructs. A robust positive correlation emerged between WTC and SI ($r = .69$, $p < .01$), affirming the pivotal role of SI in fostering communicative readiness. Similarly, a significant positive relationship was identified between WTC and FLE ($r = .64$, $p < .01$), underscoring the critical influence of positive emotional experiences in language learning contexts. Furthermore, the analysis revealed a strikingly strong correlation between SI and FLE ($r = .79$, $p < .01$), suggesting that these constructs are not only individually significant but also mutually reinforcing. These results collectively dismantle any skepticism regarding the interdependence of these variables and highlight their collective importance in shaping learners' communicative tendencies.

4.2. Results of structural equation modeling

To address the second research question—whether SI and FLE significantly predict WTC—an SEM approach was employed. This methodological choice was driven by the need to rigorously test the predictive validity of the proposed model. To evaluate the model's compatibility with the data, multiple fit indices were scrutinized against established thresholds. The χ^2/df ratio, a fundamental measure of model fit, was required to fall below 3, while indices such as the Goodness-of-Fit Index (GFI), Comparative Fit Index (CFI), Normed Fit Index (NFI), and Tucker-Lewis Index (TLI) were expected to exceed .90. Additionally, the Root Mean Square Error of Approximation (RMSEA) was constrained to values less than .08 (Collier, 2020). The results of these goodness-of-fit indices, detailed in Table 3, provide a decisive affirmation of the model's adequacy. Any deviation from these stringent criteria would have undermined the study's claims; however, the robust alignment of the data with the proposed model reinforces the argument that SI and FLE are indeed significant predictors of WTC. This empirical validation not only

strengthens the theoretical foundation of the study but also challenges prior assumptions that overlooked the predictive power of these constructs in language learning contexts.

Table 3. Goodness-of-fit indices

	χ^2/df	GFI	CFI	NFI	TLI	RMSEA
Acceptable fit	<3	>.90	>.90	>.90	>.90	<.08
Model	1.45	.97	.99	.97	.99	.04

As demonstrated in Table 3, the goodness-of-fit indices for the proposed model consistently meet or exceed the established thresholds, affirming its robust alignment with the empirical data. The χ^2/df ratio, which is a critical indicator of model fit, remains below the stringent threshold of 3, while other indices such as the GFI, CFI, NFI, and TLI all surpass the recommended value of .90 (Kline, 2023; Tseng & Schmitt, 2008; Rappaport et al., 2020). Furthermore, the RMSEA falls well below the acceptable limit of .08. These findings collectively substantiate the validity of the hypothesized model, leaving little room for contention regarding its adequacy. Consequently, it is not merely suggestive but rather conclusively evident that the proposed model is empirically supported and theoretically sound, offering a reliable framework for understanding the interplay between the constructs under investigation.

Figure 1 presents the structural model, which visually encapsulates the intricate interrelationships among EFL learners' SI, FLE, and WTC. This diagrammatic representation serves as a compelling testament to the hypothesized pathways, delineating not only the directionality of these relationships but also their statistical significance. By mapping these connections, the model elucidates how SI and FLE converge to predict WTC, thereby reinforcing the theoretical underpinnings of the study. Far from being a mere illustrative tool, Figure 1 stands as a pivotal analytical artifact, offering both clarity and empirical grounding to the complex dynamics explored in this research.

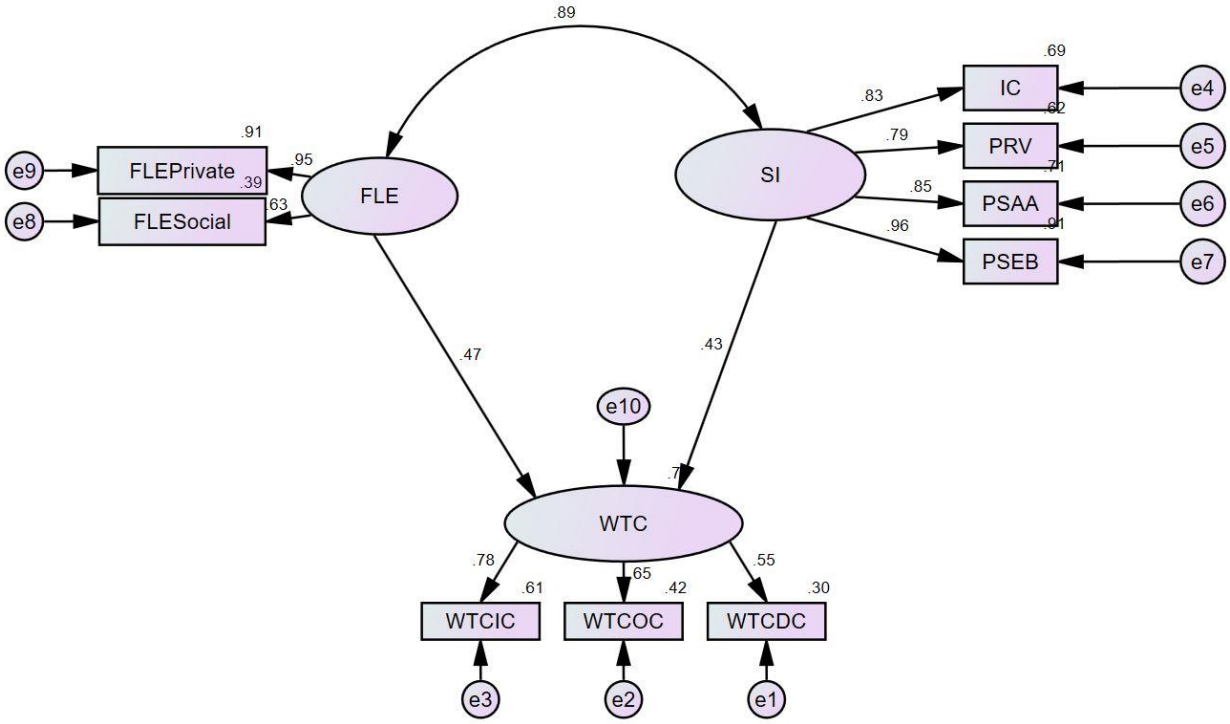

Figure 1. Interrelationships among SI, FLE, and WTC in EFL learners

The examination of standardized estimates was undertaken to rigorously interrogate the causal relationships between the independent variables—SI and FLE—and the dependent variable, WTC. As depicted in Figure 1, the findings present a compelling case for the predictive power of both SI ($\beta = .43$, $p < .05$) and FLE ($\beta = .47$, $p < .05$) on learners' WTC. These significant beta coefficients not only affirm the direct influence of these constructs but also highlight their substantial explanatory capacity within the model. Furthermore, the analysis unveils a strikingly robust positive correlation between FLE and SI ($\beta = .89$, $p < .05$), underscoring the deeply intertwined nature of these variables. This interdependence challenges any reductionist interpretation of their roles and emphasizes the necessity of considering their combined impact on communicative readiness. Collectively, these results refute skepticism regarding the causal dynamics at play, reinforcing the argument that SI and FLE are indispensable predictors in shaping learners' willingness to engage in communication.

5. Discussion

The robust positive correlation identified between SI and WTC ($r = .69$, $p < .01$) represents a decisive departure from prior studies that have either trivialized or misunderstood the behavioral impact of SI. For instance, Babazadeh et al. (2018) reported only moderate relationships between SI and language learning strategies, confining their analysis to cognitive processes such as vocabulary acquisition and strategy use. While their work was foundational, it failed to explore SI's direct influence on tangible behavioral outcomes like WTC. In stark contrast, the beta coefficient of .43 for SI's predictive power on WTC in this study establishes SI as a critical catalyst for communicative readiness—a finding that fundamentally challenges Zohar's (2005) narrow conceptualization of SI as a mechanism for existential reflection.

Critics might argue that emphasizing SI risks privileging metaphysical dimensions over pragmatic skills essential for language learning. However, the SEM employed here provides irrefutable evidence that SI's influence is neither abstract nor tangential—it is statistically significant and practically consequential. This challenges Hassan's (2009) assertion that SI fosters learning environments by focusing solely on awareness and analytical thinking, suggesting instead that its impact extends far beyond introspection to tangible behavioral outcomes. By integrating existential contemplation with observable communicative behaviors, this study reframes SI as a holistic construct that operates at the intersection of the spiritual, cognitive, and behavioral realms.

When compared to Azadi et al.'s (2022) mixed-methods approach to measuring SI, which relied heavily on self-report questionnaires, the current study's use of SEM offers a more rigorous and comprehensive validation of SI's predictive power. The reliance on standardized estimates and goodness-of-fit indices ensures that the hypothesized relationships are empirically validated rather than assumed, setting a new benchmark for methodological rigor in this field. Unlike Azadi et al., who focused narrowly on correlations between SI, gender, and educational level, this study reveals the causal pathways through which SI directly influences WTC, offering a far more nuanced understanding of its role in language learning.

The strong correlation between FLE and WTC ($r = .64$, $p < .01$) corroborates Dewaele and MacIntyre's (2014) assertion that positive emotions play a critical role in language learning. However, the marginal deviation from normality observed in FLE ($p = 0.045$) raises serious questions about the reliability of these findings. While the proximity of the p-value to the threshold justifies cautious acceptance, it also highlights potential vulnerabilities in the measurement of FLE—a limitation echoed in Pan and Zhang's (2023) longitudinal study, which found extraversion to significantly influence FLE over time. If FLE is indeed contingent upon personality traits like extraversion, then its purported universality as a predictor of WTC becomes questionable.

Moreover, the beta coefficient of .47 for FLE's predictive power on WTC suggests that enjoyment alone cannot fully account for communicative readiness. This finding directly challenges Alrabai's (2024) overly optimistic portrayal of FLE as a mediator between motivation and L2 WTC. While FLE undoubtedly fosters engagement, this study reveals that it operates in tandem with—and often subordinate to—SI. Thus, rather than viewing FLE as a standalone driver of WTC, it should be understood as part of a synergistic system where SI acts as the primary catalyst. This nuanced perspective undermines prior studies that have disproportionately emphasized emotional factors without accounting for deeper cognitive and existential dimensions.

In stark contrast to Ma et al.'s (2024) examination of teacher-student interactions and Wang and Wang's (2025) focus on emotion regulation, which overlooked the profound interconnectedness of SI and FLE, this study presents a unified theoretical framework that captures the reciprocal reinforcement between these constructs. Such an advancement challenges fragmented models of second-language acquisition and offers a more holistic explanation of communicative readiness. Prior studies, such as Dewaele et al.'s (2018) exclusive focus on learner-related factors, fail to account for the broader ecological context in which WTC unfolds. This study addresses these gaps head-on, demonstrating that FLE's impact is inseparable from the influence of SI.

Perhaps the most striking revelation of this study is the exceptionally high correlation between SI and FLE ($r = .79$, $p < .01$), which dwarfs the modest associations reported by Azizi and Zamaniyan (2013) between SI and vocabulary learning strategies. This finding not only validates King's (2009) multidimensional framework for SI—which includes recognition of the transcendent and expansion of consciousness—but also extends it by demonstrating how SI directly enhances positive emotional experiences in language learning. Critics may contend that this relationship reflects circular reasoning, given that SI inherently involves cultivating meaning and purpose, which could naturally lead to enjoyment. However, the SEM analysis provides empirical validation that this connection is neither tautological nor coincidental; it is a causal mechanism worthy of further exploration.

This synergistic relationship poses a formidable challenge to traditional models of second-language acquisition, which have historically prioritized cognitive and affective factors in isolation. For instance, Dewaele et al.'s (2018) exclusive focus on learner-related factors neglects the broader ecological context in which WTC unfolds. Similarly, Tavakoli and Davoudi's (2017) survey-based approach, which identified WTC dimensions without exploring their predictive validity, falls short of capturing the dynamic interplay between SI, FLE, and WTC. By addressing these gaps through rigorous statistical methods, including SEM and Pearson correlation analysis, this study establishes a clear causal pathway linking SI and FLE to WTC, offering a more comprehensive and empirically grounded understanding of communicative readiness.

While previous studies have made valuable contributions, they are marred by methodological shortcomings that undermine their generalizability. For example, Azadi et al.'s (2022) reliance on a single questionnaire to measure SI limits the depth of insight into its multifaceted nature. Similarly, Dewaele et al.'s (2018) exclusive focus on learner-related factors neglects

the broader ecological context in which WTC unfolds. This study addresses these gaps through rigorous statistical methods, including SEM and Pearson correlation analysis, ensuring that the hypothesized relationships are empirically validated rather than assumed.

Furthermore, the inclusion of standardized estimates and goodness-of-fit indices sets a new benchmark for methodological rigor in this field. Unlike Tavakoli and Davoudi's (2017) survey-based approach, which identified WTC dimensions without exploring their predictive validity, this study establishes a clear causal pathway linking SI and FLE to WTC. Such advancements not only enrich existing theories but also pave the way for future investigations into the mediating and moderating factors that shape these relationships.

In conclusion, this study makes an irrefutable case for the centrality of SI and FLE in fostering WTC among EFL learners. By challenging reductionist interpretations and fragmented models prevalent in prior literature, it advances a holistic understanding of communicative readiness as a product of intertwined cognitive, emotional, and existential processes. While acknowledging the slight deviation from normality in FLE, the overall findings remain robust and defensible, setting a new standard for empirical research in this domain.

6. Conclusion

This study set out to investigate the intricate relationships between SI, FLE, and WTC among EFL learners, employing rigorous statistical methodologies such as SEM and Pearson correlation analysis. The findings revealed that SI and FLE are not only significantly correlated with WTC but also operate synergistically to shape communicative readiness. Specifically, SI emerged as a primary catalyst for WTC, with a beta coefficient of .43, underscoring its pivotal role in fostering communicative behaviors. FLE, while influential with a beta coefficient of .47, was found to operate in tandem with SI rather than independently, highlighting the necessity of integrating emotional and existential dimensions in language learning. Furthermore, the exceptionally high correlation between SI and FLE ($r = .79$, $p < .01$) demonstrated their deeply intertwined nature, challenging fragmented models in prior research and offering a unified theoretical framework for understanding these dynamics. These results collectively dismantle reductionist interpretations of communicative readiness, presenting it instead as a product of intertwined cognitive, emotional, and existential processes.

The implications of this study extend to both theoretical advancements and practical applications in second-language acquisition. Theoretically, the findings challenge prevailing assumptions by reframing SI as a holistic construct that bridges existential reflection and tangible behavioral outcomes, moving beyond Zohar's (2005) narrow conceptualization of SI as purely metaphysical. This reconceptualization enriches existing frameworks, such as King's (2009) multidimensional model, by demonstrating how SI directly enhances positive emotional experiences like FLE. Practically, educators are encouraged to design curricula that integrate mindfulness, meaning-making, and emotional well-being into language learning environments. By fostering SI, educators can create transformative learning experiences that not only enhance students' communicative readiness but also cultivate deeper existential awareness. Similarly, the findings highlight the importance of nurturing FLE, though with the caveat that it must be embedded within a broader ecological context that includes SI. These insights call for a paradigm shift in how communicative readiness is conceptualized, moving beyond traditional models that prioritize either cognitive or affective factors in isolation.

While this study makes significant contributions to the field, it is not without limitations. One notable limitation is the marginal deviation from normality observed in the FLE variable ($p = 0.045$), which raises questions about the reliability of self-reported measures. This finding underscores the need for more robust instruments to capture FLE, potentially incorporating observational or experimental methods to complement self-reports. Additionally, the cross-sectional design of the study limits the ability to establish causal relationships over time. While SEM provides strong evidence for predictive validity, reciprocal influences between SI, FLE, and WTC cannot be ruled out, necessitating longitudinal designs to track the evolution of these constructs throughout the language learning process. Furthermore, the sample consisted exclusively of EFL learners from a specific cultural and educational context, which may limit the generalizability of the findings.

To address the limitations identified in this study and build upon its findings, several avenues for future research are proposed. First, longitudinal studies should be conducted to examine how SI, FLE, and WTC evolve, particularly during critical phases of language learning such as initial exposure, intermediate proficiency, and advanced fluency. Such studies could provide valuable insights into the temporal dynamics of these constructs and their reciprocal influences. Second, experimental designs could further explore the causal mechanisms underlying the relationships identified here. For example, interventions aimed at enhancing SI or FLE could be implemented to measure their direct impact on WTC, offering empirical validation of the hypothesized pathways. Third, qualitative approaches, such as interviews or focus groups, could provide richer insights into learners' subjective experiences of SI and FLE, uncovering nuances that quantitative methods might overlook. Finally, expanding the sample to include learners from diverse cultural, linguistic, and educational backgrounds would enhance the generalizability of these findings, offering a more comprehensive understanding of how SI, FLE, and WTC interact across different contexts.

7. Acknowledgements

The authors extend their heartfelt gratitude to the participants, teachers, and colleagues who contributed to this study. Special thanks to the Department of Education in Kerman Province for their support and ethical approval.

8. References

Abdollahzadeh, H., Kashmiri, M., & Arabameri, F. (2008). SI questionnaire. *Tehran: Azmoon Yar Pouya Institute*.

Aghaei, H., Behjat, F., & Rostampour, M. (2014). Investigating the relationship between Iranian high school female students' SI, language proficiency, and self-esteem. *International Journal of Language and Linguistics*, *2*(6–1), 12–19. https://doi.org/10.11648/j.ijll.s.2014020601.14

Akkas, F., Tekin, I., & Aydin, S. (2022). A Correlational study on enjoyment and anxiety in a foreign language learning context. *International LET-IN Conference*.

Alrabai, F. (2024). Modeling the relationship between classroom emotions, motivation, and learner WTC in EFL: applying a holistic approach of positive psychology in SLA research. *Journal of Multilingual and Multicultural Development*, *45*(7), 2465–2483. https://doi.org/10.1080/01434632.2022.2053138

Amram, Y., & Dryer, C. (2008). The integrated SI scale (ISIS): Development and preliminary validation. *116th Annual Conference of the American Psychological Association*, *1*, 46.

Azadi, M., Maftoon, P., & Alemi, M. (2022). Developing and validating an EFL learners' SI inventory: A mixed-methods study. *Language and Translation*, *12*(4), 87–106.

Azizi, M., & Zamaniyan, M. (2013). The relationship between SI and vocabulary learning strategies in EFL learners. *Theory and Practice in Language Studies*, *3*(5). https://doi.org/10.4304/tpls.3.5.852-858

Babazadeh, Z., Sotoudehnama, E., & Nafissi, Z. (2018). The relationship between SI, multiple intelligences, and language learning strategies. *Journal of English Language Teaching and Learning*, *10*(21), 205–222.

Benevene, P., De Stasio, S., & Fiorilli, C. (2020). Well-being of school teachers in their work environment. *Frontiers in Psychology*, *11*. https://doi.org/10.3389/fpsyg.2020.01239

Botes, E., Dewaele, J.-M., & Greiff, S. (2022). Taking stock: A meta-analysis of the effects of FLE. *Studies in Second Language Learning and Teaching*, *12*(2), 205–232. https://doi.org/10.14746/ssllt.2022.12.2.3

Burgoon, J. K. (1976). The unwillingness-to-communicate scale: Development and validation. *Communication Monographs*, *43*(1), 60–69. https://doi.org/10.1080/03637757609375916

Cao, Y. K. (2014). A sociocognitive perspective on second language classroom WTC. *TESOL Quarterly*, *48*(4), 789–814. https://doi.org/10.1002/tesq.155

Collier, J. E. (2020). *Applied SEM using AMOS*. Routledge. https://doi.org/10.4324/9781003018414

Derakhshan, A., Dewaele, J.-M., & Azari Noughabi, M. (2022). Modeling the contribution of resilience, well-being, and L2 grit to foreign language teaching enjoyment among Iranian English language teachers. *System*, *109*, 1–11. https://doi.org/10.1016/j.system.2022.102890

Derakhshan, A., Solhi, M., & Azari Noughabi, M. (2023). An investigation into the association between student-perceived affective teacher variables and students' L2-Grit. *Journal of Multilingual and Multicultural Development*, *46*(3), 798–814. https://doi.org/10.1080/01434632.2023.2212644

Dewaele, J.-M., Chen, X., Padilla, A. M., & Lake, J. (2019). The flowering of positive psychology in foreign language teaching and acquisition research. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02128

Dewaele, J.-M., & Dewaele, L. (2018). Learner-internal and learner-external predictors of WTC in the FL classroom. *Journal of the European Second Language Association*, *2*(1), 24. https://doi.org/10.22599/jesla.37

Dewaele, J.-M., & Macintyre, P. (2014). The two faces of Janus? anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, *4*, 237–274. https://doi.org/10.14746/ssllt.2014.4.2.5

Dewaele, J.-M., Witney, J., Saito, K., & Dewaele, L. (2018). FLE and anxiety: The effect of teacher and learner variables. *Language Teaching Research*, *22*(6), 676–697. https://doi.org/10.1177/1362168817692161

Dornyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge. https://doi.org/10.4324/9781315779553

Emmons, R. A. (2000). Is spirituality an intelligence? Motivation, cognition, and the psychology of ultimate concern. *International Journal for the Psychology of Religion*, *10*(1), 3–26. https://doi.org/10.1207/S15327582IJPR1001_2

Estaji, M., & Pourmostafa, P. (2020). The mediating role of SI and teaching experience in perceived leadership styles of EFL teachers: A structural equation modelling analysis. *Journal of Higher Education Policy And Leadership Studies*, *1*(1), 82–106. https://doi.org/10.29252/johepal.1.1.82

Fallah, N. (2014). WTC in English, communication self-confidence, motivation, shyness and teacher immediacy among Iranian English-major undergraduates: A SEM approach. *Learning and Individual Differences*, *30*, 140–147. https://doi.org/10.1016/j.lindif.2013.12.006

Field, A. (2024). *Discovering statistics using IBM SPSS statistics* (Sixth Edit). Sage Publications Limited.

Fredrickson, B. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. *American Psychologist*, *56*(3), 218–226. https://doi.org/10.1037/0003-066X.56.3.218

Fredrickson, B. (2003). The value of positive emotions. *American Scientist*, *91*(4), 330. https://doi.org/10.1511/2003.4.330

Fredrickson, B. (2006). The broaden-and-build theory of positive emotions. In *A Life Worth Living* (pp. 85–103). Oxford University PressNew York, NY. https://doi.org/10.1093/oso/9780195176797.003.0006

Gardner, H. (1983). *The theory of multiple intelligences*. Heinemann.

Hassan, A. (2009). Emotional and SIs as a basis for evaluating the national philosophy of education achievement. *Research Journal of International Studies*, *12*, 59–66.

King, D. B. (2009). Rethinking claims of SI: A definition, model, and measure. *Unpublished Master's Thesis, Trent University, Peterborough, Ontario, Canada*.

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.

Kruk, M. (2022). Dynamicity of perceived WTC, motivation, boredom and anxiety in Second Life: the case of two advanced learners of English. *Computer Assisted Language Learning*, *35*(1–2), 190–216. https://doi.org/10.1080/09588221.2019.1677722

Lee, J. S., & Hsieh, J. (2019). Affective variables and WTC of EFL learners in in-class, out-of-class, and digital contexts. *System*, *82*, 63–73. https://doi.org/10.1016/j.system.2019.03.002

Ma, L., Liu, L., & Liu, J. (2024). Mediation of enjoyment and burnout in the link between teacher-student relationships and foreign language achievement and its gender differences. *Innovation in Language Learning and Teaching*, 1–15. https://doi.org/10.1080/17501229.2024.2418353

MacIntyre, P. D. (1994). Variables underlying WTC: A causal analysis. *Communication Research Reports*, *11*(2), 135–142. https://doi.org/10.1080/08824099409359951

MacIntyre, P. D., Clement, R., Dornyei, Z., & Noels, K. A. (1998). Conceptualizing WTC in an L2: A situational model of L2 confidence and affiliation. *The Modern Language Journal*, *82*(4), 545–562. https://doi.org/10.1111/j.1540-4781.1998.tb05543.x

MacIntyre, P. D., & Vincze, L. (2017). Positive and negative emotions underlie motivation for L2 learning. *Studies in Second Language Learning and Teaching*, *7*(1), 61–88. https://doi.org/10.14746/ssllt.2017.7.1.4

MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching*, *2*(2), 193. https://doi.org/10.14746/ssllt.2012.2.2.4

MacIntyre, P., Gregersen, T., & Mercer, S. (2019). Setting an agenda for positive psychology in SLA: Theory, practice, and research. *The Modern Language Journal*, *103*(1), 262–274. https://doi.org/10.1111/modl.12544

McCroskey, J. C. (1992). Reliability and validity of the WTC scale. *Communication Quarterly*, *40*(1), 16–25. https://doi.org/10.1080/01463379209369817

McCroskey, J. C., & Richmond, V. P. (1990). WTC: A cognitive view. *Journal of Social Behavior and Personality*, *5*(2), 19.

Mercer, S., & MacIntyre, P. D. (2014). Introducing positive psychology to SLA. *Studies in Second Language Learning and Teaching*, *4*(2), 153–172. https://doi.org/10.14746/ssllt.2014.4.2.2

Nugroho, A. (2021). Learners' WTC in a foreign language: The role of informal digital learning of English. *ELE Reviews: English Language Education Reviews*, *1*(2), 125–133. https://doi.org/10.22515/ele-reviews.v1i2.4396

Pan, C., & Zhang, X. (2023). A longitudinal study of foreign language anxiety and enjoyment. *Language Teaching Research*, *27*(6), 1552–1575. https://doi.org/10.1177/1362168821993341

Peng, J.-E. (2007). WTC in an L2 and integrative motivation among college students in an intensive English language program in China. *University of Sydney Papers in TESOL*, *2*(1), 33–59.

Rappaport, L. M., Amstadter, A. B., & Neale, M. C. (2020). Model fit estimation for multilevel structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, *27*(2), 318-329. https://doi.org/10.1080/10705511.2019.1620109

Sato, M., & Dussuel Lam, C. (2021). Metacognitive instruction with young learners: A case of WTC, L2 use, and metacognition of oral communication. *Language Teaching Research*, *25*(6), 899–921. https://doi.org/10.1177/13621688211004639

Tavakoli, E., & Davoudi, M. (2017). WTC orally: The case of Iranian EFL learners. *Journal of Psycholinguistic Research*, *46*(6), 1509–1527. https://doi.org/10.1007/s10936-017-9504-0

Tseng, W. T., & Schmitt, N. (2008). Toward a model of motivated vocabulary learning: A structural equation modeling approach. *Language Learning*, *58*(2), 357-400. https://doi.org/10.1111/j.1467-9922.2008.00444.x

Vaughan, F. (2002). What is SI? *Journal of Humanistic Psychology*, *42*(2), 16–33. https://doi.org/10.1177/0022167802422003

Wang, Y., Derakhshan, A., & Zhang, L. J. (2021). Researching and practicing positive psychology in second/foreign language learning and teaching: The past, current status, and future directions. *Frontiers in Psychology*, *12*, 1–10. https://doi.org/10.3389/fpsyg.2021.731721

Wang, Y., & Wang, S. (2025). Boredom, enjoyment, and emotion regulation among Chinese students majoring in dual foreign languages: a Q method study. *Innovation in Language Learning and Teaching*, 1–18. https://doi.org/10.1080/17501229.2025.2450667

Wigglesworth, C. (2003). *SI: What is it? How can we measure it? Why would businesses care*? Conscious Pursuits.

Yang, L., & Lin, Y. (2024). To talk or to remain silent? FLE and perceived classroom climate as drivers of change in Chinese EFL learners' WTC: a latent growth curve analysis. *Innovation in Language Learning and Teaching*, 1–18. https://doi.org/10.1080/17501229.2024.2391376

Yashima, T., MacIntyre, P. D., & Ikeda, M. (2018). Situated WTC in an L2: Interplay of individual characteristics and context. *Language Teaching Research*, *22*(1), 115–137. https://doi.org/10.1177/1362168816657851

Yetkin, R., & ÖZER, Z. (2022). Age, gender, and anxiety as antecedents of WTC: Turkish EFL context. *Acuity: Journal of English Language Pedagogy, Literature and Culture*, *7*(2), 195–205. https://doi.org/10.35974/acuity.v7i2.2800

Yu, H., Li, H., & Gou, X. (2011). The personality-based variables and their correlations underlying WTC. *Asian Social Science*, *7*(3). https://doi.org/10.5539/ass.v7n3p253

Yu, X., & Ma, J. (2024). Modeling the predictive effect of enjoyment on WTC in a foreign language: the chain mediating role of growth mindset and grit. *Journal of Multilingual and Multicultural Development*, 1–18. https://doi.org/10.1080/01434632.2023.2300351

Zarrinabadi, N., & Pawlak, M. (Eds.). (2021). *New perspectives on WTC in a second language*. Springer International Publishing. https://doi.org/10.1007/978-3-030-67634-6

Zhang, Q., Song, Y., & Zhao, C. (2024). Foreign language enjoyment and willingness to communicate: The mediating roles of communication confidence and motivation. *System*, *125*, 103346. https://dx.doi.org/10.2139/ssrn.4539192

Zhao, J., Guan, K., Feng, Y., & Liu, Z. (2024). The effect of foreign language enjoyment on willingness to communicate among Chinese EFL students: conscientiousness as a mediator. *Frontiers in Education*, *9*, 1473649. https://doi.org/10.3389/feduc.2024.1473649

Zohar, D. (2005). Spiritually intelligent leadership. *Leader to Leader*, *2005*(38), 45–51. https://doi.org/10.1002/ltl.153

Zohar, D. (2012). *SQ SI: The ultimate intelligence* (1st ed.). Bloomsbury Publishing.

Zohar, D., & Marshall, I. N. (2001). *Spiritual intelligence: The ultimate intelligence*. A&C Black**.**

Zohar, D., & Marshall, I. (2004). *Spiritual capital: Wealth we can live by*. Berrett-Koehler Publishers.

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 121-132

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.8527.1055

Examining Strategies and Supportive Moves in Requestive Speech Acts: A Comparative Study of 'Top Notch' and 'Summit' English Series

Mohamad Reza Farangi[1*]
Zahra Kheradmand[2]

[*1]Assistant professor of Applied Linguistics, Department of Applied Humanities, Kashmar Higher Education Institute, Kashmar, Iran
[2]Ph.D. Candidate in Neuroscience, Department of Psychology, Faculty of Arts, University of British Columbia, Vancouver, Canada

ARTICLE HISTORY

Received: 16 December 2024
Revised: 05 January 2025
Accepted: 10 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: dr.farangi@kashmar.ac.ir

ABSTRACT

This study investigated the discourse features of requestive speech acts in 10 textbooks from the Top Notch and Summit series to assess their potential for developing learners' pragmatic competence. Requests were analyzed in terms of directness strategies, head acts, and supportive discourse moves, drawing on the frameworks of Blum-Kulka et al. (1989) and Farch and Kasper (1989). The findings revealed a bias towards direct request strategies in both series, likely influenced by their utilitarian discourse structure. While Top Notch provided a wider range of directness levels, Summit, designed for advanced learners, offered limited exposure to diverse request strategies. Internal supportive moves, particularly the use of questions, were more prevalent than external moves in both series. These findings suggest that while the series contribute to language learning, they may not fully equip learners with the pragmatic competence necessary for effective intercultural communication. Further research is needed to explore the impact of explicit instruction and implicit learning on the acquisition of request strategies.

KEYWORDS: Textbook evaluation; Requestive speech acts; Supportive discourse moves; Request strategies; Directness level

1. Introduction

In teaching English, incorporating the study of speech acts and their governing principles in classroom instruction offers opportunities, benefits, and difficulties. Teachers need adequate expertise to introduce these concepts and help students grasp their pragmatic dimensions while acknowledging that they can be both advantageous and demanding (Darong, 2024; Sarbandi et al., 2017). Students with limited proficiency or minimal real-world exposure to the language might find speech acts challenging. Additionally, since conventions vary significantly across languages, cultures, and social groups—even within the same group—learners must adapt to new, culturally specific forms in a foreign language (Farangi & Nejadghanbar, 2024). When interpreting meaning in a second language, learners often draw on their own cultural and linguistic experiences (Christison, 2018). To fully comprehend speech acts, one must move beyond mere fluency and explore how utterances function and how speakers use patterns to interact effectively. Thus, understanding the socio-cultural norms surrounding language use and the ways speakers convey meaning is both beneficial and essential for effective communication.

Searle (1969) categorized illocutionary acts into five types: representatives, directives, expressives, commissives, and declarations. Among these, *requests* fall under the category of directives and are considered face-threatening acts. Brown and Levinson (1987) describe requests as "*an attempt to get [the] hearer to do an act which [the] speaker wants [the] hearer to do, and*

Cite this article: Farangi, M. R. and Kheradmand, Z. (2025). Examining strategies and supportive moves in requestive speech acts: A comparative study of 'Top Notch' and 'Summit' English series. *Applied Linguistics Inquiry*, *3*(1), 121-132. doi: 10.22077/ali.2025.8527.1055

which it is not obvious that [the] hearer will do in the normal course of events or of [the] hearer's own accord" (p. 66). Thus, requests inherently challenge the interpersonal equilibrium between speakers and hearers, necessitating careful attention to levels of directness and politeness. As Blum-Kulka (1991) observes, requests are a representative way of speaking within a specific culture, reflecting broader social norms and conventions. Effective communication requires learners to understand not only the linguistic forms but also the sociocultural nuances that govern the use of requests. These include levels of directness, politeness strategies, the force of the requests, and methods for mitigating their illocutionary impact.

Reviewing various theoretical models that have been proposed to discuss request strategies and supportive moves, the current study aimed to conduct a thorough analysis of the content of the *Top Notch* and *Summit* series, which are widely used textbooks in Iranian EFL contexts. These textbooks play a significant role in shaping learners' pragmatic competence, particularly in areas such as making requests, which are essential for effective communication (Rashidi & Meihami, 2016). The primary objective of this study was to investigate the main request strategies and supportive moves presented in the requests used throughout these materials. By doing so, the study sought to determine whether these textbooks adequately expose learners to a diverse range of strategies and discourse moves for expressing requests. This examination is crucial to assess if the textbooks possess the potential to equip learners with the necessary skills to navigate different social and cultural contexts when making requests, thereby enhancing their communicative competence. Furthermore, the study aimed to identify any gaps or limitations in the presentation of request strategies, providing insights for material developers and educators to improve the pedagogical effectiveness of these resources.

This research is particularly significant given the increasing emphasis on the integration of pragmatic competence in language instruction. Incorporating the teaching of speech acts in EFL classrooms is crucial as it provides teachers with effective methods to strengthen students' conversational abilities and refine instructional techniques (Darong, 2024). By embedding speech acts into classroom activities, learners are encouraged to engage more actively and communicate with one another through realistic dialogues, such as making polite requests, giving compliments, or offering apologies (Pérez-Hernández, 2020). This method not only aids in understanding the nuances of language use but also boosts student engagement and motivation. Through structured conversational practice, students develop stronger pragmatic skills, allowing them to use English more appropriately and naturally in various contexts. Overall, focusing on speech acts in EFL instruction enriches the learning experience by combining language practice with cultural awareness, leading to more competent and meaningful communication (Bataineh et al., 2017). Requests, as a key directive speech act, are integral to achieving communicative goals in various social and professional settings. Therefore, equipping learners with the ability to navigate the nuances of requests, including appropriate strategies and mitigating discourse moves, is essential for their overall communicative success. The *Top Notch* and *Summit* series, as widely used instructional materials, have the potential to shape learners' pragmatic awareness and communicative skills, making their content worthy of detailed examination.

The choice of the *Top Notch* and *Summit* series for analysis is justified by their widespread adoption and influence in English language education globally, making them key instruments in shaping learners' pragmatic awareness and communicative competence. These textbooks are designed for intermediate to advanced learners and are frequently used in both academic and professional settings, where the development of effective communication skills is critical. As such, their content plays a significant role in how learners perceive and practice speech acts, including requests, apologies, and other pragmatic functions. A detailed examination of these materials is therefore essential to ensure that they adequately prepare learners for real-world interactions by providing comprehensive coverage of pragmatic strategies and cultural nuances (Rattanawong & Thongrin, 2022).

Moreover, the *Top Notch* and *Summit* series are notable for their integration of authentic materials and task-based activities, which aim to bridge the gap between classroom learning and practical application. However, given the complexity of pragmatic competence, it is important to evaluate whether these textbooks effectively address the diversity of request strategies and supportive moves necessary for successful communication across cultures. By scrutinizing the presentation of requestive speech acts in these widely-used materials, we can assess their adequacy in fostering learners' ability to navigate varying social and cultural contexts (Farangi & Zabbah, 2023). This analysis not only highlights areas for improvement but also contributes to the broader goal of enhancing pedagogical practices in language education (Farangi & Naami, 2024; Zhang & Su, 2021). Thus, the selection of these textbooks serves as a critical case study for understanding how instructional materials influence pragmatic development.

2. Literature review

Given the focus of this study on the *Top Notch* and *Summit* series, a review of prior research examining these textbooks' different dimensions is pertinent. Several studies have explored the linguistic and pragmatic elements of these series, offering insights into their strengths and limitations. Soozandefar and Sahragard (2011) investigated the pragmatic aspects of conversations in *Top Notch* by analyzing language functions and speech acts. They applied Halliday's framework for language functions and Searle's taxonomy of speech acts to conversations from the *Top Notch Fundamental* textbooks. Their findings indicated an unequal distribution of language functions—such as instrumental, interactional, and imaginative—and speech acts, including representatives, directives, and expressives. The researchers concluded that the conversations were pragmatically inadequate for fostering real-life communication due to their limited representation of essential functions and acts. Kafi et al. (2013) examined the impact of *Top Notch* and *Interchange* series on Iranian EFL learners' cultural attitudes. They found that these series were effective in influencing learners' perspectives on cultural issues. Notably, learners with higher English proficiency exhibited greater cultural shifts compared to those with lower proficiency levels, suggesting that advanced learners were more susceptible to the cultural content embedded in these textbooks.

Tavakol and Sayadian (2014) compared *Top Notch* and *Introductory Interchange* in terms of their potential to enhance communicative and pragmatic competence. While both series were praised for fostering speaking and writing skills through authentic materials, *Top Notch* stood out for its superior audio-visual aids and its emphasis on listening, pronunciation, intonation, and rhythm. However, *Interchange* was noted for incorporating more collaborative activities, such as pair and group work, to enhance learners' speaking abilities. Despite their differences, both series were recommended for their contributions to communicative competence. Askaripour (2014) evaluated the learning objectives of the *Top Notch* series using Bloom's revised taxonomy. The analysis revealed an unequal distribution of objectives, with a predominant focus on lower-order thinking skills (understanding, applying, and remembering) at the expense of higher-order skills (analyzing, evaluating, and creating). This imbalance underscored the need for textbooks to promote metacognitive and critical thinking skills, especially as learners' progress to higher proficiency levels.

Ahmadi Safa et al. (2015) explored EFL teachers' and learners' attitudes towards the potential of *Top Notch* to foster intercultural awareness and competence. Their study assessed dimensions such as cultural knowledge, intercultural skills, and cross-cultural understanding. While both groups appreciated the series' cultural content, teachers—due to their greater experience and language proficiency—expressed higher expectations for the series to comprehensively address intercultural competence. Homayounzadeh and Sahragard (2015) assessed the ability of *Top Notch* and *Summit* to foster intercultural communicative competence through an analysis of sociocultural identities (cultural, ethnic, gender, and personal). Their findings highlighted the strengths of *Top Notch* in representing diverse global cultures, while *Summit* emphasized American and European cultural norms to acquaint advanced learners with Western communicative practices. Both series displayed a Utilitarian discourse structure, characterized by solidarity and deference, though *Summit* also incorporated elements of individualism to reflect Western ideologies.

Regarding international studies, Boxer and Pickering (1995) examined seven ELT textbooks structured around the teaching of language functions to identify issues in how they present speech acts. The analysis centers on the specific sequence of complaining and commiserating to illustrate the disparity between spontaneous, naturally occurring speech and the contrived examples created by textbook developers based on their native-speaker intuition. One key issue highlighted is that the intuitive assumptions about how speech acts are realized often diverge significantly from actual patterns found in natural speech. Additionally, the study points out that crucial information regarding the underlying social strategies of speech acts is frequently neglected. To highlight this contrast, a sample lesson on complaining and commiserating based on authentic, spontaneous speech is provided, which serves to underscore the differences between the lessons presented in the surveyed textbooks and those grounded in real-world communication.

Kawashima (2022) explored how ELT textbooks provide learners with opportunities to develop pragmatic knowledge, specifically focusing on common speech acts like giving advice, making offers, extending invitations, and making requests. To analyze how these speech acts are presented, the study examined five international beginner-level textbooks and seven Japanese high school ELT textbooks. The results showed that international textbooks typically incorporate these speech acts to simulate everyday conversational interactions. In contrast, Japanese high school textbooks primarily use them to explain grammar rules or, in some cases, to practice strategies for softening language. Both types of textbooks, however, share limitations, such as a restricted variety of situational contexts and insufficient pragmatic options for learners. Based on these findings, the study emphasizes the importance of teachers adapting textbook materials by modifying their internal features to better align with pragmatic teaching goals.

Pramono and Kurniawan (2020) investigated the frequency of pragmatic content related to the speech acts of thanking and apologizing in two Indonesian ELT textbooks: one prescribed by the Ministry of National Education and one commercially published. The analysis of the pragmatic content in the textbooks, titled *When English Rings the Bell* and *English on Sky*, was conducted using Martinez's Framework for the speech acts of thanking and apologizing. The findings revealed that both the prescribed and commercial textbooks adequately covered instances of these speech acts. However, despite the existence of multiple strategies for performing thanking and apologizing, neither textbook included a diverse range of these strategies. As a result, it is recommended that language teachers incorporate additional authentic language materials to better develop learners' pragmatic competence.

In a 2023 study by Wilson, an examination of the speech acts targeted for instruction in ELT (English Language Teaching) textbooks commonly used in Hong Kong was conducted through relational content analysis. The goal was to identify which speech acts are frequently featured or entirely absent in these materials, allowing for comparison with findings from similar studies conducted in other regions. This is significant from a pedagogical standpoint, as teachers must ensure that textbooks equip students with the ability to use a broad spectrum of speech acts effectively. Additionally, textbooks should provide clear examples and explanations to help learners understand the appropriate contexts and methods for using these speech acts. In this research, the presentation of speech acts was analyzed to assess the inclusion of pragmalinguistic (e.g., indirect speech acts) and sociopragmatic (e.g., handling high social distance) information. The results indicated a tendency to prioritize certain speech acts over others, with some being completely omitted. Furthermore, there was a notable deficiency in the provision of pragmalinguistic and sociopragmatic details. These findings suggest that both textbook authors and educators should reconsider how such materials are developed and selected, highlighting areas for potential improvement. Overall, this study offers valuable insights into the portrayal of speech acts in Hong Kong ELT textbooks, with important implications for textbook design and pragmatic instruction in English language teaching.

A study conducted by Ren and Han (2016) revealed that ELT textbooks used in China inadequately represented pragmatic content. The researchers noted a scarcity of contextual information, with the presentation of speech acts often relying on the intuition of the textbook writers rather than empirical data. They suggested that these materials could benefit from additional metapragmatic information, reasoning, and details about intralingual variation. Following this research, subsequent studies by Pérez-Hernández

(2019) and Ton Nu and Murray (2020) also identified a consistent lack of pragmatic information in ELT textbooks. These findings highlight a broader trend of insufficient coverage of pragmatic aspects, underscoring the need for improvement in how such content is addressed in language teaching materials.

Despite these extensive evaluations, the pragmatic aspects of requestive speech acts in *Top Notch* and *Summit* remain unexplored. Building on frameworks proposed by Blum-Kulka, Farch, and Kasper (1989), and House and Kasper (1989), this study aimed to address this gap by analyzing request strategies and supportive moves in these textbooks. Specifically, the study examined the following:

1. The distribution of request strategies in the *Top Notch* and *Summit* series.
2. The distribution of internal and external supportive moves in requestive speech acts within these textbooks.
3. The co-occurrence matrix of internal and external supportive moves with each request strategy.

The analysis focuses on the textbooks' potential to familiarize learners with various strategies and discourse moves for expressing requests. By addressing these questions, the study seeks to contribute to the broader understanding of the pragmatic content of these widely used textbooks and their role in enhancing learners' communicative competence. Therefore, this study attempted to provide answers to the following questions:

1. What is the distribution of request strategies in Top Notch and Summit series?
2. What is the distribution of internal and external supportive moves used in requestive speech acts in Top Notch and Summit series?

3. Methodology

3.1. Theoretical framework

Building on this foundation, Blum-Kulka et al. (1989) conceptualized requests as speech acts consisting of obligatory *head acts*—the core of the request—and optional elements known as *supportive moves*. Supportive moves are designed to modify and mitigate the illocutionary force of the head acts. Farch and Kasper (1989) classify supportive moves into two categories: (a) internal supportive moves and (b) external supportive moves.

Internal supportive moves involve lexical and syntactic modifications applied directly within the request head acts to soften the degree of directness. Lexical modifications include mitigators, such as hedging expressions, and mental verbs, which temper the assertiveness of the request. Syntactic modifications involve structural adjustments, such as the use of conditional statements or interrogative forms, to render requests less imposing. External supportive moves, by contrast, are articulated in separate utterances, either preceding or following the head act. These include strategies such as providing reasons, using preparatory statements, offering disarming remarks, presenting alternatives, employing precursors, and integrating positive politeness strategies. For example, a preparatory statement might signal the forthcoming request, while a disarming remark could alleviate potential resistance. Understanding and applying these strategies are critical for learners aiming to achieve pragmatic competence. Table 1 (not included here but suggested for inclusion in the full text) could provide illustrative examples of both internal and external supportive moves, enhancing learners' awareness of how requests are formulated and mitigated in different contexts. This framework underscores the importance of embedding pragmatic instruction within textbook content to ensure learners acquire not only linguistic knowledge but also the sociopragmatic skills necessary for effective interaction in a second language.

Table 1. Examples of both internal and external supportive moves

A. Internal Supportive Moves
1. Lexical
a) Use of mitigators: E.g., would you please open the window?
b) Use of mental verbs: E.g., I think proofreading my article would take you less than half an hour.
2. Syntactic
a) Use of Conditionals: E.g., if you are free now, please go to the post office and send the letter for me.
b) Use of Questions: E.g., Could you wash the dishes?
B. External Supportive Moves
1. Providing Reasons: E.g., because I do not have enough money I cannot buy it now. Can you keep it for me till tomorrow?
2. Use of preparators: E.g., you know that our company is very careful about the quality of the materials. The rice that was imported this week was not good enough and we returned it back. Can you give us one week extension so that we provide you with the best-quality rice?
3. Use of disarmers: E.g., I have heard that you are so knowledgeable and proficient. Would you please take a look at my project and give me your comments?
4. Use of precursors/alerters: E.g., excuse me sir! May you show me the post office?

5. Suggesting alternatives: E.g., I am so sorry John, I have forgotten my wallet. You pay for me now, tomorrow I will pay for you. Agree?
6. Use of positive politeness strategies: E.g., here, we have a notebook in which our visitors write their names and their emails in order to keep in touch with them. However, it is up to you, do as you wish.

Blum-Kulka et al. (1989) classified requests based on different levels of directness to direct, conventionally indirect, and non-conventionally indirect requests. The illocutionary force in direct requests is perceived directly thorough linguistic elements. In Conventionally indirect requests, the utterance's meaning is indicated through linguistic norms already used in a specific context. In non-conventionally indirect requests, the force of the utterance is calculated based on the interactions between the actions and contextual references. Different categories of request strategies along with their examples are presented below.

3.1.1. Direct level

1. Mood derivable: An Utterance in which the grammatical structure or a specific manner which is used indicate the most direct requests (E.g., Hand in your homework.).

2. Performatives: An utterances in which the illocution is clearly stated (E.g., I tell you to clean your room before noon.)

3. Hedged performatives: An utterance in which the force of the utterance is mitigated by hedges (E.g., Would you please send the file to me?).

4. Obligations: An utterance in which the hearer is forced to do an act (E.g., You should tell the manager about your project.)

5. Want statements: An utterance in which the speaker's desires and wants are stated (E.g., I want you to call me as soon as possible.).

3.1.2. Conventionally indirect level

6. Suggestory: An utterance in which the illocution is stated through a clear suggestion (E.g., Why not trying the new program?)

7. Query-preparatory: An utterance in which both the preliminary conditions such as, ability and eagerness, and the real requests are stated (E.g., Would you mind collecting the papers?).

3.1.3. Non-conventionally indirect level

8. Strong hints: An utterance in which the indirect request contains some reference to the elements of the act and is considered an implicit suggestion (E.g., It is very hot in this room.).

9. Mild hints: An utterance in which there is no reference to the request's elements and the request needs to be interpreted within the context (E.g., She has called you five times, but you have been busy.)

3.2. Materials

The investigation of strategies and discourse moves in requestive speech acts was carried out using textbooks selected from the *Top Notch* and *Summit* series, designed by Saslow et al. (2011). The *Top Notch* series comprises six textbooks, each tailored to different proficiency levels, aiming to engage learners in English interactions and communications effectively. Each textbook follows a uniform structure, including the following sections:

1. Topic Preview – An introduction to the theme of the unit.
2. Conversation Practice – Dialogues designed for practicing natural language use.
3. Grammar Practice – Conversations focusing on grammatical structures.
4. Vocabulary Section – Key vocabulary items for the unit.
5. Reading Comprehension – A passage aimed at improving reading skills.
6. Listening and Pronunciation Activities – Exercises targeting auditory and phonological skills.

The *Summit* series, 3rd Edition, is a two-level course designed for high-intermediate to advanced learners, focusing on communication skills. It prepares confident, culturally-aware individuals to manage social, travel, and professional situations in English. This course can be used after completing the intermediate level of any communicative series, such as Top Notch. It comprises 4 textbooks each containing new design features updated content aimed at enhancing language skills through various engaging resources. Conversation Activator videos are designed to improve communicative competence, while Discussion Activator videos encourage both the quality and quantity of expression. Additionally, the Test-taking Skills Booster, along with extra challenge exercises, supports students in mastering the reading and listening sections of standardized tests. There is also a

significant increase in practice opportunities for grammar, reading, listening, and writing. Each unit in the *Summit* textbooks begins with Topic Preview and Sound Bites, where learners discuss real-life issues and practice natural conversations. This is followed by:

1. Conversation Snapshot – Introducing idiomatic expressions.
2. Grammar Snapshot – Presenting advanced grammatical structures.
3. Reading and Writing Sections – Activities aimed at improving literacy skills.
4. Listening and Pronunciation Activities – Integrated exercises to enhance auditory comprehension and phonological awareness.

Notably, the *Summit* textbooks place a greater emphasis on listening and writing skills compared to the *Top Notch* series, aligning with the needs of advanced learners. Both series are corpus-based, developed using the Longman Corpus Network, which incorporates data from the Longman Spoken Corpus and the Learners' Corpus of Common Learner Errors. This foundation ensures that the materials reflect authentic language use and address frequent learner errors, making them pedagogically robust.

In summary, the *Top Notch* and *Summit* series are designed to provide comprehensive language learning experiences, with a clear progression from beginner to advanced levels. Their structure and content are instrumental in facilitating the investigation of requestive speech acts, offering rich material for analyzing strategies and discourse moves in pragmatic contexts.

3.3. Data collection and analysis procedures

The primary objective of this content analysis of the *Top Notch* and *Summit* series was to evaluate these textbooks' potential for familiarizing EFL learners with the request speech act and equipping them with diverse strategies to perform it effectively. Given the centrality of pragmatic competence in communicative language teaching, this study aimed to assess how well these series address the complex nature of requests and their associated strategies. The findings are critical in determining the extent to which these materials provide learners with tools to navigate real-world communication, where the ability to formulate and interpret requests is essential for maintaining interpersonal relationships and achieving communicative goals. The analysis process involved systematically identifying and extracting all instances of requests embedded in the texts and compiling them into a dedicated dataset for further scrutiny. This systematic approach ensured a thorough examination of the materials while maintaining the integrity of the data. The analysis was conducted at two interrelated levels to provide a detailed understanding of the requestive speech acts and their pragmatic underpinnings.

To address the research questions evaluating the potential of the *Top Notch* and *Summit* series in introducing the request speech act to EFL learners, six *Top Notch* textbooks spanning six proficiency levels and four *Summit* textbooks designed for advanced learners were subjected to a detailed content analysis. In our analysis, we examined all instances of requestive moves across the entire scope of the textbooks, encompassing speaking, listening, reading, and writing tasks. This comprehensive approach allowed us to evaluate the various contexts in which requestive moves are presented and practiced, ensuring a well-rounded understanding of their usage and application throughout the materials. Since the series were designed for different proficiency levels, we focused on comparing books within comparable proficiency bands (e.g., beginner-to-intermediate or intermediate-to-advanced). For instance, if one series targeted beginners while the other targeted intermediates, we compared the higher-level books of the first series with the lower-level books of the second series. This alignment ensured that the comparison was made between materials intended for learners with similar linguistic capabilities.

Descriptive statistics were utilized to analyze the frequency and distribution of these elements across the series. The quantitative approach enabled a detailed comparison of the prevalence of various strategies and supportive moves, shedding light on the pragmatic emphasis—or lack thereof—in the instructional materials. To ensure the reliability and validity of our analysis, we employed a rigorous inter-coder reliability process. Two independent coders categorized all instances of requestive moves identified in the textbooks using the framework developed by Blum-Kulka et al. (1989). After the initial coding, discrepancies between the two coders were discussed and resolved through consensus, ensuring alignment with the established categories of direct, conventionally indirect, and non-conventionally indirect strategies. The inter-coder agreement was calculated using Cohen's Kappa coefficient, which measures the level of agreement beyond chance. The resulting Kappa value was 0.85, indicating a high level of reliability.

4. Results

The findings are presented in two sections: first, the analysis of the *Top Notch* series, followed by the *Summit* series. As shown in Table 2, the *Top Notch* series contained a total of 153 instances of requests across the six textbooks. A breakdown of the requests by levels of directness revealed that direct requests were the most frequent, comprising 89 instances (58.2%). Non-conventionally indirect requests were the second most prevalent category, accounting for 55 instances (35.9%). In contrast, conventionally indirect requests constituted only 9 instances (5.9%), representing a notably smaller proportion of the total requests.

This unequal distribution suggests that the textbooks prioritize introducing learners to direct forms of requests over more nuanced or indirect forms. While direct requests can facilitate learners' initial understanding of pragmatic functions, the limited representation of conventionally indirect requests may hinder the development of more contextually appropriate and

culturally sensitive communication skills, especially in formal or high-stakes situations where indirectness is often preferred. In contrast to the *Top Notch* series, the *Summit* series included significantly fewer instances of requests, with a total of only 20 examples across the four textbooks. Among these, direct requests were again the most common, accounting for 13 instances (65%). Conventionally indirect requests appeared 6 times (30%), while non-conventionally indirect requests were represented by just a single instance (5%).

The smaller corpus of requests in the *Summit* series, coupled with the predominance of direct strategies, suggests a similar emphasis on straightforward communication. Despite the advanced proficiency level of the intended learners, the textbooks do not appear to provide a balanced exposure to the full range of pragmatic strategies needed for making requests in diverse social contexts.

Table 1. Frequency of requests in each level of directness in Top Notch and Summit series

Level of directness	Top Notch F	Summit F
Direct	89	13
Conventionally indirect	9	6
Non-conventionally indirect	55	1
Total	153	20

While it is true that the *Top Notch* series included two additional books compared to the *Summit* series in our analysis, the descriptive data clearly indicates a substantial discrepancy in the frequency of request speech acts between the two series (*Top Notch* : 153 instances vs. *Summit* : 20 instances). This significant difference cannot be attributed solely to the disparity in the number of books within each series. Instead, it reflects a more fundamental distinction in how the authors of each series approach the treatment of this pragmatic speech act. This is further supported by the fact that our analysis carefully compared books from both series that were closely matched in terms of learners' proficiency levels, as outlined in the methodology section. Despite this alignment, the frequency and variety of request strategies presented in the *Top Notch* series far exceed those in the *Summit* series. This suggests that the authors of *Top Notch* place greater emphasis on the development of pragmatic competence through the explicit instruction of request strategies and supportive moves, whereas the *Summit* series appears to allocate less attention to this aspect of communicative ability. Given the rigorous matching of proficiency levels and the striking contrast in the presentation of request speech acts, we argue that these comparisons are valid and reliable.

The findings reveal that the request speech act is underrepresented across the four levels of the *Summit* series, indicating a significant gap in the content designed for advanced learners. This omission limits the learners' opportunities to practice and refine their understanding of request strategies at advanced levels, which is crucial for developing nuanced pragmatic competence. However, this approach does not adequately address the diverse learning pathways of students. Many learners may transition to the *Summit* series after studying elementary and intermediate levels using different textbooks. For these students, a lack of content on request speech acts in the *Summit* series means missing an essential opportunity to develop this skill. Furthermore, even for learners who have studied the *Top Notch* series, regular practice and reinforcement of request strategies are essential at advanced levels. Advanced learners are expected to navigate more complex and context-sensitive social interactions, which require a deeper understanding and mastery of indirect and nuanced request strategies.

On the other hand, as shown in table 3, within the *Top Notch* series, there was an imbalanced use of direct strategies. Hedges were overrepresented, appearing far more frequently than other strategies. Conversely, obligations and performatives were underrepresented, with only a limited number of instances provided. This lack of balance restricts learners' exposure to the variety of direct strategies available for making requests, which may impede their ability to choose appropriate strategies in different contexts. The representation of conventionally indirect strategies was notably insufficient. Only two examples were identified: one **suggory** and one **query-preparatory**. This minimal exposure is inadequate to familiarize learners with these important pragmatic strategies, which are often crucial in polite and formal interactions. Such limited representation fails to equip learners with the tools needed to perform conventionally indirect requests effectively.

For non-conventionally indirect strategies, the books demonstrated some efficiency in incorporating strong hints, which were the most frequently used strategy in this category. However, the inclusion of mild hints was inadequate, providing learners with insufficient examples to understand the subtleties of using mild hints for indirect requests. As previously mentioned, the *Summit* series lacks sufficient content on request speech acts. Many strategies, such as performatives, obligations, want statements, and mild hints, were entirely absent. This omission not only limits the pragmatic richness of the textbooks but also fails to address the increasing complexity of language use required at advanced proficiency levels. Advanced learners must be able to adjust their language use to a variety of social and cultural contexts, making the absence of such content a notable shortcoming.

The findings highlight a significant need for the *Summit* series to include more comprehensive coverage of request speech acts, with a balanced representation of direct, conventionally indirect, and non-conventionally indirect strategies. Advanced learners require exposure to nuanced language use, including indirectness and politeness strategies, to navigate a wider range of communicative situations effectively. Addressing these gaps would ensure that learners who transition to the *Summit* series—whether or not they have used the *Top Notch* series—receive adequate instruction and practice in request speech acts.

Moreover, even for learners familiar with requests at the lower levels, advanced-level textbooks must provide opportunities for review and expansion of these skills, tailored to the learners' increasing proficiency and communicative needs.

Table 3. Frequency of different request strategies used in each degree of directness in Top Notch and Summit series

Direct level	**Top Notch**	**Summit**
	F	**F**
Mood derivable	15	8
Performatives	0	0
Hedged performatives	61	5
Obligations	1	0
Want statement	12	0
Total	89	13
Conventionally indirect level	**F**	**F**
Suggestory	6	1
Query-preparatory	3	5
Total	9	6
Non-conventionally indirect level	**F**	**F**
Strong hints	52	1
Mild hints	3	0
Total	55	1

In addition to the request strategies used for expressing head acts, internal and external supportive moves play a crucial role in mitigating the illocutionary force of these head acts. As shown in Table 4, the overall frequency of supportive moves in the *Top Notch* series was considerably higher than that in the *Summit* series, primarily due to the greater number of requests identified in the *Top Notch* books compared to the *Summit* series. Overall, internal supportive moves were employed more frequently than external ones to modify and reduce the illocutionary force of the head acts. However, within the *Top Notch* series, the distribution of internal supportive moves was uneven across different request types. As indicated in Table 4, expressing requests in the form of questions was the most commonly used internal move. This trend was also observed in the *Summit* series, where questions predominated among the supportive moves employed. Regarding external supportive moves, *preparators* emerged as the most frequently used strategy to soften the illocutionary force of requests. These findings suggest a consistent preference for specific internal and external strategies in both series, though their overall distribution and frequency varied significantly between the two.

Table 4. Frequency of different internal and external supportive moves in requestive speech acts found in Top Notch and Summit series

Internal Supportive Moves	**Top Notch**	**Summit**
	F	**F**
Lexical (use of mitigators)	34	5
Lexical (use of mental verbs)	13	0
Syntactic (questions)	60	15
Syntactic (conditionals)	18	0
Total	125	20
External Supportive Moves	**F**	**F**
Providing reasons	6	5
Use of preparators	5	0
Use of disarmers	1	0
Use of precursors/alerters	6	0
Suggesting alternatives	0	1
Use of positive politeness strategies	0	1
Total	18	7

In addition to investigating the frequency of request strategies and supportive moves separately, this study investigated the co-occurrence of different supportive moves with different strategies, the results of which are presented in Table 5.

Table 5. Frequency of internal and external supportive moves used with each request strategy

Supportive Moves \ Strategies	Mood derivable	Performatives	Hedged Performatives	Obligations	Want statement	suggestory	Query-preparatory	Strong hints	Mild hints
Lexical (use of mitigators)	11	-	4	-	1	-	-	21	-
Lexical (use of mental verbs)	-	-	-	-	-	-	-	12	-
Syntactic (questions)	-	-	28	-	-	5	4	12	1
Syntactic (conditionals)	-	-	3	-	-	-	1	-	-
Providing reasons	-	-	6	-	-	-	1	-	-
Use of preparators	1		1	-	-	-	-	1	1
Use of disarmers	-	-	1	-	-	-	-	-	-
Use of precursors/alerters	-	-	1	-	1	-	-	4	-
Suggesting alternatives	1	-	-	-	-	-	-	-	-
Use of positive politeness strategies	-	-	-	-	-	-	-	1	-

As shown in Table 5, requests in the form of hedge performatives mitigated by questions occurred most frequently, with a total of 28 instances. Following this, strong hints combined with mitigators were the second most common group of requests, appearing 21 times. Strong hints were also equally mitigated by mental verbs and questions, with each combination occurring 12 times. Additionally, the illocutionary force of requests expressed through mood derivables was primarily softened by the use of mitigators, with this combination occurring 11 times. However, it is important to note that some requests were made solely using request strategies, without the inclusion of any internal or external moves.

5. Discussion

The aim of this study was to investigate the pragmatic aspects of requestive speech acts in the Top Notch and Summit series, two widely used ELT textbooks. Specifically, the research sought to analyze the distribution of request strategies and supportive moves within these textbooks to determine their potential to familiarize learners with various strategies and discourse moves for expressing requests. Drawing on frameworks proposed by Blum-Kulka et al. (1989), the study addressed three key questions: (1) What is the distribution of request strategies in the Top Notch and Summit series? (2) What is the distribution of internal and external supportive moves used in requestive speech acts? By answering these questions, the study aimed to contribute to a deeper understanding of the pragmatic content of these textbooks and their role in enhancing learners' communicative competence.

The findings revealed significant disparities in the representation of request strategies across the Top Notch and Summit series. In the Top Notch series, a total of 153 instances of requests were identified, with direct requests being the most frequent (89 instances, 58.2%), followed by non-conventionally indirect requests (55 instances, 35.9%), and conventionally indirect requests (9 instances, 5.9%). In contrast, the Summit series included only 20 instances of requests, with direct requests again dominating (13 instances, 65%), followed by conventionally indirect requests (6 instances, 30%) and a single instance of non-conventionally indirect requests (5%). This unequal distribution suggests that both series prioritize direct forms of requests over more nuanced or indirect forms, potentially limiting learners' exposure to culturally sensitive communication strategies. Additionally, the Summit series lacked comprehensive coverage of request speech acts, particularly for advanced learners who require nuanced language skills. The analysis of internal and external supportive moves further highlighted an uneven distribution, with questions being the most frequently used internal move and preparators being the predominant external move.

Similar to the results of this study, Soozandefar and Sahragard (2011) found an unequal distribution of language functions and speech acts in the Top-Notch series, emphasizing its pragmatic inadequacy for real-life communication. Similarly, our analysis revealed a significant imbalance in the representation of request strategies across the Top Notch and Summit series. The higher the level of the books, the lower the distribution of speech acts were. Direct requests dominated, comprising 58.2% of instances in the Top-Notch series and 65% in the Summit series. Conventionally indirect and non-conventionally indirect requests were underrepresented, accounting for only 5.9% and 35.9%, respectively, in the Top-Notch series and even less so in the Summit series (30% and 5%, respectively). This skewed focus on direct requests may hinder learners' ability to develop nuanced communication skills necessary for navigating diverse social contexts, echoing Soozandefar and Sahragard's critique of the series' limited pragmatic scope.

Ren and Han (2016), Pérez-Hernández (2019), and Ton Nu and Murray (2020) further underscored the insufficient coverage of pragmatic aspects in ELT textbooks globally. The minimal inclusion of conventionally indirect requests in the studied textbooks reflects this broader trend, suggesting that these materials fail to provide learners with adequate exposure to culturally sensitive and contextually appropriate language use. However, Homayounzadeh and Sahragard (2015) identified the Utilitarian discourse system as a defining feature of both the Top Notch and Summit series. This system prioritizes egalitarian relationships, free expression, and straightforward communication, minimizing hierarchical social structures. Our findings confirm this influence, as characters in the textbooks are predominantly portrayed as having equal social power, leading to the predominance

of direct requests. While such an approach fosters clarity and equality, it neglects the complexities of real-world interactions where indirectness and politeness strategies are often required. This limitation resonates with Boxer and Pickering's (1995) observation that textbook examples frequently diverge from authentic, spontaneous speech patterns, failing to capture the underlying social strategies of speech acts.

Moreover, Kafi, Ashraf, and Motallebzadeh (2013) noted that advanced learners are more susceptible to cultural content embedded in textbooks, highlighting the importance of addressing their specific needs. However, the Summit series, designed for advanced learners, included significantly fewer instances of requests compared to the Top-Notch series (20 vs. 153). Furthermore, many essential strategies—such as performatives, obligations, want statements, and mild hints—were entirely absent. This omission limits opportunities for advanced learners to refine their understanding of request strategies, despite their increasing need for nuanced language use in complex communicative situations. In addition, Askaripour (2014) criticized the Top Notch series for focusing predominantly on lower-order thinking skills at the expense of higher-order ones. Similarly, the lack of comprehensive coverage of request strategies in the Summit series suggests a failure to promote critical thinking and metacognitive awareness among advanced learners. Addressing this gap would ensure that learners can analyze, evaluate, and adapt their language use to suit varying social and cultural contexts.

On the other hand, Tavakol and Sayadian (2014) praised the Top Notch series for its emphasis on listening, pronunciation, intonation, and rhythm but noted areas for improvement in fostering pragmatic competence. Our findings reveal that internal supportive moves, particularly questions, were employed more frequently than external ones to mitigate the illocutionary force of requests. While this preference for specific strategies provides some consistency, the uneven distribution of supportive moves across different request types restricts learners' exposure to the full range of pragmatic tools available. For instance, preparators emerged as the most frequently used external move, yet other strategies like disarmers and precursors/alerters were rarely or never employed. This imbalance mirrors the overall pattern of underrepresentation observed in the textbooks' pragmatic content. Besides**,** Kawashima (2022) compared international and Japanese high school ELT textbooks, finding that both types shared limitations in situational variety and pragmatic options. The Top Notch and Summit series similarly exhibit restricted contextual diversity, offering learners limited opportunities to practice request strategies in varied settings. Wilson's (2023) study of Hong Kong ELT textbooks highlighted deficiencies in pragmalinguistic and sociopragmatic details, reinforcing the need for textbooks to provide clear explanations and examples of speech act usage. The co-occurrence matrix of internal and external supportive moves with request strategies in our study underscores the importance of addressing these gaps to enhance learners' pragmatic competence effectively.

The findings suggest that while the Top Notch and Summit series introduce learners to basic request strategies, they fall short in preparing them for real-world communicative demands. In line with recommendations from previous studies, instructors should supplement these materials with authentic language resources and activities to bridge the gap between textbook content and practical application. Additionally, future editions of these textbooks could incorporate a balanced representation of direct, conventionally indirect, and non-conventionally indirect requests, along with diverse supportive moves, to better equip learners with the skills needed for effective communication.

6. Conclusion

This study investigated the presentation of requestive speech acts in the *Top Notch* and *Summit* English language learning series. By analyzing the frequency and types of request strategies employed, as well as the use of supportive discourse moves, we sought to assess the extent to which these series equip learners with the necessary pragmatic competence to engage in effective intercultural communication. Our findings indicate that while both series offer valuable exposure to English language use, they exhibit certain limitations in their treatment of requestive speech acts. *Top Notch*, designed for lower-level learners, provides more opportunities for request practice, particularly through the use of direct requests. However, it could benefit from a more balanced approach, incorporating a wider range of strategies, including indirect requests and politeness markers. *Summit*, aimed at advanced learners, presents fewer opportunities for request practice, potentially neglecting the development of pragmatic competence in this area. In conclusion, the present study contributes to the growing body of research on the pragmatic content of ELT textbooks by providing a detailed analysis of request strategies and supportive moves in the Top Notch and Summit series. By linking these findings to existing literature, we emphasize the urgent need for improvements in how textbooks address pragmatic aspects of language learning, particularly for advanced learners requiring nuanced and context-sensitive communication skills.

To enhance the pedagogical effectiveness of language learning series like Top Notch and Summit, material developers should consider several key strategies. First, diversifying request strategies by incorporating direct, indirect, and conventionally indirect requests can better cater to various social and cultural contexts (Farangi & Khojastemehr, 2024). Emphasizing supportive moves such as hedges, softeners, and politeness markers is crucial for mitigating face-threatening acts and fostering positive interactions. Additionally, integrating contextual factors like power relations, social distance, and cultural norms into the curriculum can help learners understand the nuances of appropriate communication. Explicit instruction on these strategies, supported by clear explanations, examples, and practice exercises, is essential. Moreover, encouraging learner production through ample opportunities to practice written and spoken requests in diverse scenarios will solidify their understanding and application.

Despite the valuable insights provided by this study, certain limitations must be acknowledged. The corpus used may not fully represent the breadth of request strategies and supportive moves within these series, and the analysis was conducted from a Western cultural perspective, potentially overlooking cultural variations. Furthermore, the focus on explicit instruction

might neglect the role of implicit learning through]authentic language exposure. To address these limitations, future research could explore longitudinal studies tracking learners' development over time, examine learner perceptions of request strategies, and conduct cross-cultural comparisons of language learning materials. Investigating the role of input enhancement techniques and technology-enhanced learning tools, such as CALL and VR, could also provide new avenues for improving pragmatic competence (Farangi et al., 2024). By addressing these gaps, we can enhance our understanding of requestive speech act acquisition and refine language learning materials and practices accordingly.

7. References

Ahmadi Safa, M., Moradi, M., & Hamzavi, R. (2015). Iranian EFL teachers and learners perspective on potentiality of Top Notch series for intercultural competence development. *Iranian Journal of Language Teaching Research 3* (2), 47-66.

Akbari Kelishadi, A., & Sharifzadeh, A. (2013). An evaluation of Top Notch series. *International Journal of Language Learning and Applied Linguistics World, 4* (4), 60-73.

Askaripour, S. A. A. (2014). A textbook evaluation of new version (2nd edition) of "Top Notch English series. *English for Specific Purposes World, 15* (44), 1-30.

Austin, J. (1962). How to do things with words. In A. Jaworski & N. Coupland (Eds.), *the discourse reader (*pp. 63-75). New York: Routledge.

Bataineh, R. F., Al-Qeyam, F. R., & Smadi, O. M. (2017). Does form-focused instruction really make a difference? Potential effectiveness in Jordanian EFL learners' linguistic and pragmatic knowledge acquisition. *Asian-Pacific Journal of Second and Foreign Language Education*, *2*, 1-11. https://doi.org/10.1186/s40862-017-0040-0

Blum-Kulka, S. (1991). Interlanguage Pragmatics: The Case of Requests. In R. Phillipson, E. Kellerman, L. Salinker, M. Sharwood Smith & M. Swain (eds.), *foreign/second language pedagogy research* (pp. 255–272). Clevedon and Philadelphia: Multilingual Matters.

Blum-Kulka, S., House, J., & Kasper, G. (1989). Investigating cross-cultural pragmatics: An introductory overview. *Cross-cultural pragmatics: Requests and apologies*, *31*, 1-34.

Blum-Kulka, S., House, J., & Kasper, G. (Eds.). (1989). *Cross-cultural pragmatics: Requests and apologies*. Norwood: Ablex Publishing.

Boxer, D., & Pickering, L. (1995). Problems in the presentation of speech acts in ELT materials: The case of complaints. https://doi.org/10.1093/elt/49.1.44

Brown, P., & Levinson, S.C. (1987). *Politeness: Some universals in language usage*. Cambridge: Cambridge University Press.

Christison, M. (2018). Speech act theory and teaching speaking. *The TESOL Encyclopedia of English Language Teaching*, 1-6. https://doi.org/10.1002/9781118784235.eelt0712

Darong, H. C. (2024, June). Teaching speech acts in EFL context. In *Frontiers in Education* (Vol. 9, p. 1423498). Frontiers Media SA. https://doi.org/10.3389/feduc.2024.1423498

Farangi, M. R., & Khojastemehr, M. (2024). Iranian applied linguists (mis) conceptions of ethical issues in research: a mixed-methods study. *Journal of Academic Ethics*, *22*(2), 359-376.

Farangi, M. R., & Naami, A. (2024). A study of early multilingualism in Iran: The effects of mother tongue literacy, socio-economic status and foreground TV exposure. *Acta Psychologica*, *243*, 104167.

Farangi, M. R., & Nejadghanbar, H. (2024). Investigating questionable research practices among Iranian applied linguists: Prevalence, severity, and the role of artificial intelligence tools. *System*, *125*, 103427.

Farangi, M. R., Rashidi, N., & Sanjarani, A. (2024). A Functional Analysis of EFL Classroom Talk: The Case of Experience. *MEXTESOL Journal*, *48*(1), n1.

Farangi, M., & Zabbah, I. (2023). Intelligent scoring in an English reading comprehension course using artificial neural networks and neuro-fuzzy systems. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *42*(4), 1-21.

Farch, C., & Kasper, G. (1989). Internal and external modification in interlanguage request

Homayounzadeh, M., & Sahragard, R. (2015). An Intercultural Approach to Textbook Evaluation: A Case of Top Notch and Summit Series. *International Journal of Applied Linguistics & English Literature, 4* (1), 198-208. https://doi.org/10.7575/aiac.ijalel.v.4n.1p.198

Kafi, Z., Ashraf, H., & Motallebzadeh, K. (2013). English textbooks and cultural attitudes of Iranian EFL learners: investigating a relationship. *International Journal of Language Learning and Applied Linguistics World, 4* (2), 25-36.

Kawashima, C. (2022). Learners' Opportunities to Practice Speech Acts in English Language Teaching Textbooks. https://doi.org/10.37546/JALTPCP2021-03

Pramono, S. A., & Kurniawan, E. (2020). a Micro Analysis Study of Pragmatic Content in Indonesian Elt Textbooks. *Jurnal Pendidikan Bahasa*, *9*(1), 92-102. https://doi.org/10.31571/bahasa.v9i1.1673

Pérez-Hernández, L. (2019). From Research to the Textbook: Assessing Speech Acts Representation in Course Book Series for Students of English as an L2. *Spanish Journal of Applied Linguistics, 32*(1), 248–276. https://doi.org/10.1075/resla.17024.per

Pérez-Hernández, L. (2020). *Speech Acts in English: From Research to Instruction and Textbook Development*. Cambridge University Press.

Rashidi, N., & Meihami, H. (2016). Hidden curriculum: An analysis of cultural content of the ELT textbooks in inner, outer, and expanding circle countries. *Cogent Education*, *3*(1), 1212455. https://doi.org/10.1080/2331186X.2016.1212455

Rattanawong, A., & Thongrin, S. (2022). *An evaluation of cultural contents in ELT textbooks used by university students in Thailand and development of an ELT textbook redesigned under EIL paradigm* (No. 308286). Thammasat University.

Ren, W., & Han, Z. (2016). The Representation of Pragmatic Knowledge in Recent ELT Textbooks. ELT Jour-nal, 70(4), 424–434. https://doi.org/10.1093/elt/ccw010

Sarbandi, F. R., Taki, G., Yousefian, P., & Farangi, M. R. (2017). The effect of physicians' gender and experience on Persian medical interactions. *Discourse and Interaction, 10*(1), 89-110.

Saslow, J. M., Ascher, A., & Rouse, J. C. (2011). *Top notch.* Second edition. Pearson Longman.

Searle, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge: Cambridge University Press.

Soozandefar, S. M. A., & Sahragard, R. (2011). A Textbook Evaluation of Speech Acts and Language Functions in Top-Notch Series. *Theory and Practice in Language Studies, 1* (12), 1831-1838.

Tavakol, M., & Sayadian, S. (2014). An evaluation of Top Notch fundamental A and B. *International Journal of Research Studies in Language Learning, 3* (2), 101-114. http://dx.doi.org/10.5861/ijrsll.2013.411

Ton Nu, A. T., & Murray, J. (2020). Pragmatic Content in EFL Textbooks: An Investigation into Vietnamese National Teaching Materials. *TESL-EJ, 24*(3), 1–28

Wilson, L. D. (2023). Investigating the coverage of speech acts in Hong Kong ELT textbooks. *Electronic Journal of Foreign Language Teaching*, *20*(1). https://doi.org/10.56040/lmdw2014

Zhang, X., & Su, X. (2021). A cross-national analysis of cultural representations in English textbooks used in China and Germany. *SN Social Sciences*, *1*(4), 91. https://doi.org/10.1007/s43545-021-00088-8

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 133-146

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.8873.1090

Predicting Teachers' Stroke, through Identity, Immunity, and TPACK: Iranian EFL Teachers in Focus

Shabnam Kadkhodaei[1]
Mohammad Reza Talebinejad[2]
Mohsen Shahrokhi[3*]

[1]Ph.D. Candidate in Applied Linguistics, Department of English Language, Shahreza Branch, Islamic Azad University, Shahreza, Iran
[2,*3]Associate Professor of Applied Linguistics, Department of English Language, Shahreza Branch, Islamic Azad University, Shahreza, Iran

ARTICLE HISTORY

Received: 12 December 2024
Revised: 21 January 2025
Accepted: 24 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: mohsen.shahrokhi@iau.ac.ir

ABSTRACT

Despite extensive research on teacher qualities, a significant knowledge gap persists regarding the interplay among psychological and professional factors influencing Iranian EFL teachers' performance, particularly concerning teacher stroke. This study aims to address this gap by examining the predictive power of teacher identity, immunity, and Technological Pedagogical Content Knowledge (TPACK) on teacher stroke, within the Iranian EFL context. Participants included 124 Iranian EFL teachers, randomly selected from various regions across Iran, who completed a set of validated questionnaires. Data were collected using the Teacher Immunity Questionnaire, TPACK Questionnaire, Teacher's Identity Questionnaire, and Teachers' Stroke Scale. Statistical analysis involved Multiple Regression Analysis to explore the predictive capacity of teacher identity, immunity, and TPACK on stroke. Results revealed a statistically significant relationship among these variables, indicating that teacher identity, immunity, and TPACK collectively explain 46% of the variance in teacher stroke. These findings underscore the importance of integrating these factors into frameworks for supporting teachers' well-being and enhancing teaching effectiveness. The implications extend to stakeholders in EFL education, suggesting targeted interventions to bolster teacher development and instructional quality.

KEYWORDS: EFL teachers; Identity; Immunity; TPACK; Stroke

1. Introduction

Professionals in the field of English as a Foreign Language (EFL) education have recently become more interested in the issue of how various teacher-related factors interact to influence teaching effectiveness and student outcomes. The role of teachers in shaping educational outcomes has long been recognized as a critical determinant of student success and the overall effectiveness of educational systems (Burroughs et al., 2019; Derakhshan et al., 2020).

Among the myriad factors influencing teaching quality, teacher-related variables such as stroke, identity, immunity, and Technological Pedagogical Content Knowledge (TPACK) have emerged as crucial components that impact both teacher well-being and instructional effectiveness. Teacher stroke, defined as any action acknowledging and valuing another person's presence (Shirai, 2006), has been shown to influence teachers' professional effectiveness, student motivation, and commitment to attending classes (Pishghadam & Khajavy, 2014; Pishghadam et al., 2021). Similarly, teacher identity, which encompasses personal characteristics, relational connections, and group affiliations, plays a significant role in professional development and teaching performance (Sedikides & Brewer, 2015; Yazan, 2023). Teacher immunity, conceptualized as the defense mechanism teachers employ to mitigate negative effects on their professional identity (Hiver, 2015), is also integral to understanding educators' self-concept and resilience

Cite this article: Kadkhodaei, S., Talebinejad, M. R. and Shahrokhi, M. (2025). Predicting teachers' stroke, through identity, immunity, and TPACK: Iranian EFL teachers in focus. *Applied Linguistics Inquiry*, *3*(1), 133-146. doi: 10.22077/ali.2025.8873.1090

(Mercer, 2018). Furthermore, TPACK represents the intersection of technological, pedagogical, and content knowledge, essential for the effective integration of technology in educational settings (Mishra & Koehler, 2006; Tseng et al., 2020).

Despite extensive research on these individual factors, there remains a significant gap in understanding how they interact within specific educational contexts, particularly among Iranian EFL teachers. Previous studies have examined teacher stroke (e.g., Afsharpour et al., (2023); Frymier et al., 2019; Pishghadam et al., 2021; Pishghadam et al., 2019; Pishghadam & Farkhondefal, 2017; Pishghadam & Khajavy, 2014; Rajabnejad et al., 2017), identity (Atai et al., 2022; Ghiasvand et al., 2023; Kalali Sani et al., 2021; Moradkhani & Ebadijalal, 2024), immunity (e.g., Azari Noughabi et al., 2024; Gooran et al., 2023; Namaziandost et al., 2023; Wang et al., 2024 to name a few), and TPACK (Najjari et al., 2021; Nessaei & Kazemi, 2024; Raygan & Moradkhani, 2022), but few have explored their interplay and collective impact on teaching practices (Kadkhodaie et al., 2023; Jiang et al., 2021). This study aims to address this gap by investigating the relationships among these variables within the Iranian EFL context, where unique challenges and opportunities exist due to limited English use outside formal educational settings (Pishghadam et al., 2019).

The rationale for selecting these variables stems from their potential to provide comprehensive insights into the multifaceted nature of teaching effectiveness. Understanding how teacher stroke, identity, immunity, and TPACK collectively influence EFL teaching can inform targeted interventions to enhance teacher well-being and instructional quality. For instance, identifying the predictive power of these factors could lead to the development of support systems that bolster teachers' professional identities and technological competencies, ultimately improving educational outcomes (Pishghadam et al., 2021).

This study holds significant implications for stakeholders in EFL education. With the increasing availability and widespread utilization of technology, educators recognize the significance of incorporating it into the curriculum and pedagogical practices. Moreover, the potential effects of technology on the professional growth and performance of EFL teachers have become particularly significant. By elucidating the complex interactions among teacher stroke, identity, immunity, and TPACK, the current research offers valuable insights for teacher training programs, educational policymakers, and practitioners. The findings could inform the design of professional development initiatives that address the specific needs of Iranian EFL teachers, thereby enhancing their teaching effectiveness and contributing to broader educational reforms (Derakhshan et al., 2020).

2. Literature review

Teaching English as a foreign language (TEFL) is a challenging profession that requires a range of competencies, including linguistic proficiency, pedagogical knowledge, and teaching skills. However, effective teaching is not solely determined by these factors, as personal as well as environmental variables can also have a big impact on teachers' well-being and teaching effectiveness (Kyriacou, 2001). In this literature review, we will explore the literature on stroke, teacher's identity, teacher's immunity, and TPACK in the context of EFL instruction, with a focus on Iranian EFL teachers.

2.1. Teacher stroke

The existing literature on teacher stroke is extensive and focuses particularly on educational settings. Teacher stroke enacted by appreciating and recognizing the students' presence can, among other things, promote interpersonal skills, motivation, learning, and stress reduction in the educational setting (Peng & Woodrow, 2010; Pierson, 2003). One of the most important techniques for studying interpersonal connections between teachers and students is Berne's (1988) Transactional Analysis (TA), which is described as "a theory of personality and a systematic psychotherapy for personal growth and personal change" (Stewart & Joines, 1987, p. 3). Particularly in educational settings, the use of TA yields more effective and transparent teacher-student exchanges (Stewart & Joines, 1987). According to Berne (1988), TA theory incorporates aspects of interactions, self-worth status, social status, temporal systems, and stroke. Any action that indicates an awareness of another person's presence is referred to as a unit of recognition, and the stroke component expresses this (Shirai, 2006). Stroke, specifically in an educational setting, refers to instructors' concern for and attention to their students (Rajabnejad et al., 2017).

Positive and negative, verbal and nonverbal, and conditional and adversative strokes are all possible. Positive strokes produce the strokee's happiness and delight in this regard, whilst negative strokes cause the strokee to feel unsatisfied (Song et al., 2022). The exchange of communication, which might take the form of one word or an extended conversation, is referred to as verbal strokes (Berne, 1988). Contrarily, non-verbal cues include actions like smiling, nodding, and caressing among others. Unconditional strokes are more concerned with what individuals are, as opposed to conditional strokes, which are related to people's behaviors (Sauter, 2017).

Since it can encourage pupils to do well and increase the likelihood that the stroked action will recur, stroke plays a crucial function in the classroom (Stewart & Joines, 1987). Researchers have taken note of the stroke concept, based on an analysis of the pertinent literature, and various studies have looked at how it interacts with other factors. The EFL students' academic achievement (Khorsand & Modarresi, 2023), motivation of students (Pishghadam & Khajavi, 2014), students' socioeconomic status (Irajzad & Shahriari, 2017), teacher burnout (Yazdanpour, 2015), teacher gender identity (Noorbakhsh et al., 2018), and teacher success (Pishghadam et al., 2019), for instance, have all been linked to teacher stroke. In the EFL context, past literature reveals an association between a teacher's stroke with L2 students' achievement (Freedman, 1993), L2 education, and dedication to classroom attendance (Namaghi, 2016; Pishghadam et al., 2021), in addition to the teacher's stroke typology conducted by Pishghadam and Farkhondehfal (2017).

While some studies emphasize the positive impact of teacher stroke on student motivation (Pishghadam & Khajavy, 2014; Rajabnejad et al., 2017), others argue that its effects may vary depending on contextual factors such as cultural norms and classroom dynamics. A better understanding of the nature of a teacher's stroke could be developed by finding the other teacher's variables that can affect the function and influence of teacher stroke. Language Teacher's immunity, among other teacher variables, that has been studied (Azari Noughabi et al., 2024; Gooran et al., 2023; Namaziandost et al., 2023) in the EFL contexts per se is a well-being teacher's variable whose association with EFL teacher's stroke is still unexplored. This study is, therefore, an attempt to cast further light on this unexplored research area.

2.2. Language teacher immunity

The immunity of language teachers is a significant factor that has a profound impact on their behavior, attitude toward their profession, and professional identity, as indicated by several studies (Haseli et al., 2018; Hiver, 2015, 2016). Essentially, the immunity of English language instructors refers to the defense mechanisms that language teachers employ to mitigate the negative effects of disruptions on their professional identity as L2 teachers (Hiver & Dörnyei, 2017). In simpler terms, teacher immunity serves as a shield for teachers, equipping them with the necessary tools to handle stress in complex situations and provide successful teaching (Hiver, 2017).

Similar to physiological immunity, the defense mechanism of language teachers manifests in two ways when faced with difficult circumstances: productive and maladaptive. When language instructors are passionate, resilient, and motivated, they are said to have productive immunity, which secures them from adverse situations. On the contrary, teachers who have maladaptive immunity experience emotional fatigue, negativity, and apathetic conditions (Hiver & Dörnyei, 2017). In Iran, Haseli et al. (2018) investigated the most common form of immunity across the English instructors' community and checked how they obtained their immunity. The study revealed that maladaptive immunity predominated among Iranian EFL teachers, but did not specify the factors that influence the different types of immunity during various stages of immunity development. Rahmati et al. (2019) examined how immunity developed among Iranian EFL instructors and found that instructors' poor confidence in themselves, low earnings, low student enthusiasm, constraints on time, and substantial parental demands were the main disruptions that sparked their immunity.

While Hiver (2017) highlights the dual nature of teacher immunity as both protective and restrictive, Noughabi et al. (2020) suggest that immunity's influence is more nuanced, particularly when mediated by variables like teacher engagement and professional identity. As language teacher immunity significantly impacts teachers' career choices and actions (Hiver, 2015), investigating its relationship with teachers' identity and stroke provides valuable insights (Hiver & Dörnyei, 2017). Despite this, the pertinent literature has not explored the factors that, along with immunity, can predict the EFL instructors' stroke. Hence, there is a need to conduct research on the predictive power of the immunity of language instructors and its relationship with their job identity (Hiver, 2016) in an EFL context.

2.3. Teacher professional identity

Recent research on teacher development has given more attention to the concept of professional identity (Moradkhani & Ebadijalal, 2024). This has led to investigations into teachers' perspectives, narratives, and biographies (Rodgers & Scott, 2008), which provide insights into the factors that influence an individual's inclination toward a particular job (Tsakissiris, 2015). According to Tsakissiris (2015), people form their identities based on their individual characteristics, relational connections, and group affiliations. As Skorikov and Vondracek (2011) maintain a person's concept of individual identity heavily depends on their sense of job or professional identity. Professional identity plays a significant role in the growth and development of a nation (Low et al., 2012). The process of developing a professional identity is thought to be influenced by a variety of situations and personal factors, and by external factors such as political, social, and cultural circumstances, as well as interactions with others, being critical in shaping a person's identity (Bressler & Rotter, 2017).

Language teacher professional identity has recently heeded much from L2 teacher researchers (De Costa & Norton, 2017). Teacher identity development is now seen as an essential component of instructor education and is a prominent area of research in the training of teachers (Tsui, 2003). According to Morgan and Clarke (2011), the identity of language teachers is a crucial aspect of the ongoing reconstruction of their expertise and abilities in the field of language instruction. Such a perspective aligns with the recent sociocultural trends in the field of SLA teaching, which aim to present instructor expertise not as a solitary array of intellectual abilities but as inherently associated with factors such as the growth and identity of instructors (Johnston et al., 2005). This approach suggests that investigating the development of teachers' identities can provide insight into how language teachers develop as experts as they transition from a student self to an instructor self.

Language education studies, in recent years, have focused on the construction and development of teacher professional identity (Atai et al., 2022; Ghiasvand et al., 2023; Kalali Sani et al., 2021) and on investigating different aspects of language instructors' job identity development in both L2 instruction and teacher preparation. Scholars have explored the connection between language teacher identity and their linguistic identities, race, and gender (Aneja, 2016; Huang, 2014; Kayi-Aydar, 2015; Rudolph, 2016; Park, 2017; Yazan, 2017; Vitanova, 2016). Moreover, they have examined how language teachers negotiate discourses in communities of practice, the impact of internships in developing identity, teachers' positioning and agency in teaching contexts, the role of emotions in identity development, and how teacher education courses affect identity negotiation (Trent, 2017; Wolff & De Costa, 2017; Yazan, 2017, 2018). While teacher professional identity is a well-established research line studied from various

perspectives as briefly reviewed above, its association with teacher stroke in the development of a conceptual model that is a major objective for the current study is still calling for further studies.

2.4. Teacher TPACK

In the field of education, TPACK is a conceptual model (Mishra & Koehler, 2006) that explains how teachers can incorporate technology into their instruction. TPACK is derived from Shulman's Pedagogical Content Knowledge (PCK) (1986), serving as its foundation. According to Shulman, a teacher's effectiveness depends on their knowledge of content and pedagogy. TPACK extends Shulman's PCK and proposes that teachers must cultivate three essential knowledge domains: technological knowledge (TK), pedagogical knowledge (PK), and content knowledge (CK). By developing these domains, teachers can effectively teach subject matter content using specific instructional methods and technology in various contexts (Bostancıoğlu & Handley, 2018).

Educators and researchers have taken an interest in the TPACK framework, with around 600 publications currently listed in the Scopus database and a number of papers in the Web of Science, as reviewed by Tseng et al. (2020). The researchers conducted a critical evaluation of TPACK studies and identified several implications for future research, including exploring the relationship between TPACK and other teacher-related variables. Numerous investigations have been carried out in this field, assessing teacher TPACK knowledge (Najjari et al., 2021; Raygan & Moradkhani, 2022), and also examining the link between instructors' TPACK, their professional development, educational ideas, and self-efficacy (Chai et al., 2013; Kavanoz et al. 2015; Liu & Kleinsasser, 2015; Nessaei & Kazemi, 2024; Mei et al., 2018; Saudelli & Ciampa, 2016; Yang, 2018). However, there is a lack of understanding of the findings in TPACK research related to other teacher-related factors such as teachers' immunity, professional identity, and stroke, particularly in an EFL context like Iran.

Previous research has highlighted the significance of the teacher's identity, immunity, and TPACK in EFL teaching effectiveness (e.g., Jiang et al., 2021; Koehler & Mishra, 2009). However, few studies have examined the relationships among these factors and their potential interactions with stroke in the Iranian EFL context. As noted by Pishghadam et al. (2018), understanding the complex interplay between different factors affecting EFL teachers in Iran is crucial for developing effective support systems and interventions that can enhance their well-being and teaching effectiveness. Therefore, the present study aims to explore the relationships between teacher stroke, identity, immunity, and TPACK in Iranian EFL teachers. Moreover, it is an attempt to examine the predictive role that the teacher identity, immunity, and TPACK probably play on the variation of stroke. Such a framework can inform the development of effective interventions and support systems that can enhance EFL teachers' well-being and teaching effectiveness in Iran. Understanding the intricate nature of EFL teaching will be made easier with the help of this research and it can provide insights for teacher training programs and educational policymakers. Therefore, the current study seeks to answer the following qustions:

1. Is there a significant association among Iranian EFL teacher stroke, Identity, Immunity, and TPACK?
2. Do Immunity, Identity, and TPACK significantly predict Iranian EFL teacher stroke?

3. Methodology

3.1. The design of the study

This study employed a quantitative, non-experimental research design to investigate the relationships among variables under investigation. The research design was structured around a correlational framework, which allowed for the examination of associations among the variables without manipulating them. Specifically, the study aimed to determine the extent to which teachers' immunity, TPACK, and identity predict their stroke, as well as to explore the interplay among these constructs. This approach enabled the creation of a robust model that not only identified significant predictors but also quantified their contributions to the variance in the dependent variable.

3.2. Participants and settings

The study involved a total of 124 EFL teachers, selected non-randomly based on their availability and their consent to have a sample of Iranian EFL instructors from various regions in Iran, ensuring geographic and institutional diversity. Among the participants, 68 were male (54.8%) and 56 were female (45.2%), reflecting a relatively balanced gender distribution. The mean age of the participants was 34.2 years (*SD* = 7.8), indicating a diverse range of teaching experience levels.

3.3. Research instruments

This study employed four instruments to collect the required data as follows.

3.3.1. Teacher immunity questionnaire

The study's quantitative data were collected using a teacher immunity questionnaire, created by Hiver (2016), which consists of two sections. Demographic information was gathered in the first section, including teaching experience, gender, and age. The

second section of the questionnaire consisted of 39 items divided into seven subscales, including coping, self-efficacy of instructor, burnout, resilience, views on instruction, willingness to change, and affectivity. These items were measured using a seven-point Likert scale, ranging from *Strongly Disagree* to *Strongly Agree*. The questionnaire displayed an overall Cronbach's alpha index of 0.85, indicating strong internal consistency (Haseli Songhori et al., 2018). This instrument was translated into Persian by Haseli Songhori et al. (2018), and inter-coder reliability was established through back-translation by two professional translators, with a reported Cohen's Kappa of 0.88, indicating high reliability.

3.3.2. TPACK questionnaire

In this study, Bagheri's (2020) TPACK tool was utilized because its validity and reliability have been established in an EFL context. The tool included 31 items that utilized a seven-point Likert scale, measuring three sources of knowledge: CK, PK, and TK. Interactions between these three sources created four additional knowledge sources: technological pedagogical knowledge (TPK), technological content knowledge (TCK), pedagogical content knowledge (PCK), and technological pedagogical content knowledge (TPACK). Additionally, the questionnaire evaluated teachers' self-assessed confidence in utilizing web-based teaching and searching for subject-specific web materials. These items were incorporated under a new category called web content knowledge (WCK). In the current study, the total Cronbach's alpha for the questionnaire was determined to be 0.86.

3.3.3. Teacher's identity questionnaire

The third survey used in the study was the teacher's identity questionnaire, adapted from Liou (2008). The questionnaire comprises two parts. The first part, Section A, included eight questions that were designed based on the components of professional identity for non-native English teachers and focused on their views regarding the social status of their job, their job commitment and evaluation, and their perceptions of their pupils' ideas about their instructions. Following the initial section, Section B included 18 questions intended to gauge respondents' views on language proficiency in different contexts, such as grammar and pronunciation, their preferences in teaching content, and whether they believe the purpose of learning English is to achieve language proficiency and accuracy or intelligibility in communication. The teacher's identity questionnaire demonstrated a reliability of 0.85.

3.3.4. Teachers' stroke scale

This questionnaire was developed and validated by Kadkhodaie et al. (2023) to measure the EFL teacher's perceived stroke practice. The scale consists of 34 items, with choices varying between *Strongly Disagree* and *Strongly Agree*, as a five-point Lickert scale. The items on the scale assess various aspects of teachers' classroom behavior and interactions with students, such as their ability to greet and motivate students, their use of disciplinary measures, their relationship with students, and their use of up-to-date and real-life teaching materials. The construct validity of the Teacher Stroke Scale was checked by Kadkhodaie et al. (2023) using two methods: exploratory factor analysis (EFA) and confirmatory factor analysis (CFA). The convergent validity was established by examining the correlations between the items on the scale and their respective factors, which were found to be significant and in the expected direction (Kadkhodaie et al., 2023). The scale demonstrated strong internal consistency, as evidenced by a Cronbach's alpha coefficient of .87, indicating that the items in the scale effectively measured a shared concept. The Teacher Stroke Scale provides an overall estimate of EFL teachers' perceived stroke practice stroke based on teachers' classroom behavior and interactions with students.

3.4. Data collection procedure

The procedure began on March 15, 2023, with circulating a consent letter describing a detailed explanation of the study's objectives, the purpose of the questionnaires, and the ethical considerations, including the confidentiality of responses and the voluntary nature of participation. Participants, contacted in person or via email and social media, were assured that the data collected would be used solely for research purposes.

The research instruments were distributed to 200 EFL teachers who had expressed their willingness to participate. These teachers were from various language schools across Iran's central, eastern, western, northern, and southern regions to ensure geographic diversity and a representative sample. The data collection spanned over a three-month period, until June 20, 2023. This timeline allowed participants sufficient time to complete the questionnaires thoughtfully and return them without feeling rushed.

Out of the 200 distributed questionnaires, 124 were successfully completed and returned, achieving a response rate of 62%. Upon receiving the completed questionnaires, the data were carefully reviewed for completeness and accuracy before statistical analysis (Creswell & Creswell, 2017).

3.5. Data analysis

The data analysis phase involved a comprehensive approach to examine the relationships, using the Statistical Package for the Social Sciences (SPSS, version 26), among the study's key variables. To achieve this, Pearson Product Moment Correlation was

used to check the associations among variables; moreover, Multiple Regression Analysis was employed for the exploration of predictive power, enabling the creation of a model that quantifies the contributions of independent variables (teacher identity, immunity, and TPACK) to the dependent variable (stroke) (Hahs-Vaughn & Lomax, 2020). The analysis aimed to determine the extent to which these factors collectively explain the variance in teacher stroke and identify their relative significance as predictors.

4. Results

4.1. The association among variables

The first research question addressed the correlation among the main variables of the study. To answer the question, Pearson Correlation Coefficient was run, after ensuring that the data was normally distributed and there was no significant outlier. As the results presented in Table 1 indicate a series of significant associations among Iranian EFL teachers' stroke, identity, immunity, and TPACK were observed. This conclusion was supported by robust evidence derived from the Pearson correlation coefficients (r), which revealed that all correlations between the main variables are positive and statistically significant ($p < .001$ for all correlations). Specifically, Teachers' Stroke demonstrated strong correlations with Identity ($r = .516$), Teachers' TPACK ($r = .546$), and Teachers' Immunity ($r = .582$). Similarly, Identity was significantly correlated with Teachers' TPACK ($r = .454$) and Teachers' Immunity ($r = .448$), while Teachers' TPACK showed a notable correlation with Teachers' Immunity ($r = .527$). The significance levels (Sig. 1-tailed) for all correlations were reported as $p < .000$, indicating that the observed associations are highly unlikely to have occurred by chance. These findings provide compelling evidence of statistically significant relationships among the variables.

Table 1. Correlations among main Variables

		Teacher Stroke	Teacher Identity	Teacher TPACK	Teacher Immunity
Pearson Correlation	Teacher Stroke	1.000	.516	.546	.582
	Teacher Identity		1.000	.454	.448
	Teacher TPACK			1.000	.527
	Teacher Immunity				1.000

Furthermore, the strength of the correlations ranges from moderate (.448 to .527) to strong (.546 to .582), suggesting meaningful and substantive associations among the variables. For instance, the strongest correlation is observed between Teachers' Stroke and Teachers' Immunity (r = .582), followed closely by the correlation between Teachers' Stroke and Teachers' TPACK (r = .546). Understanding the relationships among these variables can inform targeted interventions aimed at enhancing EFL teachers' professional development and instructional quality.

4.2. The predictive power of independent variables

To answer the second research question that was raised regarding the predictive power of teacher identity, immunity, and TPACK on stroke, Multiple Regression analysis was run. In order to guarantee the validity and reliability of the findings derived from a Multiple Regression analysis, specific assumptions described by Hahs-Vaughn and Lomax (2020) were checked in the current study. To check one of the assumptions the Cook's Distance statistic was run for each participant. Typically, values of the Cook's Distance statistic exceeding 1 are regarded as significant outliers that could exert excessive influence on the regression model, necessitating their removal. The current study did not identify any such outliers, confirming that the assumption was satisfied. Moreover, the data presented in Table 1 demonstrated a linear relationship between the independent variables and the dependent variable, thereby fulfilling the requirement of another assumption for running Multiple Regression analysis.

To verify an additional assumption, which is collinearity, before conducting the Multiple Regression analysis, it is crucial to ensure that the predictors (or independent variables) do not exhibit excessively high correlation. This can be accomplished through two methods. Firstly, by examining the correlation table provided above, it is important to observe that no correlations exceed 0.8 among the predictors. In this study, there was no concern regarding this matter, as the strongest correlation observed was r =.582. Moreover, the two collinearity diagnostic tests, namely *Tolerance* or *VIF* (Variance Inflation Factor) could also be used as the second more objective way of testing the collinearity assumption. A threshold for VIF that is often regarded as indicating more substantial collinearity among predictors is 10 (Pituch & Stevens, 2016) which would correspond to a tolerance of 0.2. Hence, in order to satisfy the collinearity assumption, the VIF test values should be significantly lower than 10, and the Tolerance test value scores should be above 0.2. These conditions were successfully met in the present study.

Regarding the subsequent assumption, which necessitated that the residuals exhibit a normal distribution, a normal P-P plot of Regression Standardized Residual was used. The results indicated that the majority of the data points closely align with the line, indicating the fulfillment of another assumption. The Durbin-Watson test was used for checking the next assumption. The Durbin-Watson statistic is utilized to verify the independence assumption of the residuals. This statistic falls within the range of 0 to 4, and to satisfy the assumption, the value should be close to 2. In this research, however, the Durbin-Watson value was quite near 2 (*Durbin-Watson* = 1.954), indicating that the assumption has been satisfied.

The *R*-value indicated the magnitude of the association between the dependent variable, which is the teacher stroke, and the collective predictor variables, such as immunity, identity, and TPACK. In the present research, the *R*-value was 0.68, denoting a robust relationship (Cohen, 1988). The *R*-square value (R^2 = .463) illustrated the proportion of variability in the dependent variable, that is accounted for by the included predictors. The combination of predictors, including immunity, identity, and TPACK, explained around 46% of the variability in Teacher Strokes. According to Cohen (1988), the impact size of the predictor set on the dependent variable, as measured by R-square, can be interpreted using the following benchmarks: $R^2 \leq .02$ (small effect); $R^2 = .13$ (medium effect); $R^2 \geq .26$ (large effect).

Table 2. ANOVA[a]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	7.332	3	2.444	34.552	.000[b]
	Residual	8.487	120	.071		
	Total	15.819	123			

a. Dependent Variable: Teacher Stroke
b. Predictors: (Constant), Teacher TPACK, Teacher Identity, Teacher Immunity

In Table 2, the F-ratio is used to evaluate the suitability of the regression model for the data. It tests whether the independent variables, including immunity, professional identity, and TPACK, predict the dependent variable, the teacher's stroke, significantly. According to the table, $F(3, 120) = 34.552$ and $P < .000$, which suggests that the regression model is well-suited for the data and that the independent variables have a significant predictive ability for the teacher's stroke.

To meet the next assumption, that is, homoscedasticity, the scatterplot of the residuals was checked. Observationally, the scatterplot did not indicate any homoscedasticity, since the residuals were not evenly scattered across the plot. Therefore, a Modified Breusch-Pagan test was run to check the assumption of homoscedasticity statistically. The results indicated that the assumption of homoscedasticity was not met. Heteroskedasticity of the data may either increase Type 1 or Type 2 error probability when making inferences regarding regression parameters (Hayes & Cai, 2007). In these cases, one potential approach to reduce the biasing effect of heteroskedasticity on test statistics and inferences is to use regression using the HC3 estimator, as displayed in the following table.

Table 3. Parameter estimates with robust standard errors

Dependent Variable: Teacher Stroke

				95% Confidence Interval	
Parameter	B	Robust Std. Error[a]	t	Lower Bound	Upper Bound
Intercept	1.678	.426	3.941	.000	.835
Identity	.313	.140	2.241	.027	.037
TPACK	.130	.047	2.789	.006	.038
Immunity	.161	.055	2.907	.004	.051

a. HC3 method

Table 3 provides information on the degree to which each individual predictor variable makes a significant contribution to the model. Looking at the Sig. column, it is revealed that the teacher identity ($p = 0.027$), immunity ($p = 0.004$), and TPACK ($P = 0.006$) as predictors of the model, could all significantly predict the teacher stroke, as the dependent variable of the study. The Beta column in the table presents the unstandardized beta coefficients (B values) for the model. The coefficients provide insights into the associations between the dependent variable (teacher stroke) and all the predictor variables (teacher identity, immunity, and TPACK). Essentially, these *B* values offer an understanding of the impact each predictor has on the outcome while holding the impacts of other predictors constant. It follows that for every one-unit increase in the teacher identity mean score, the teacher stroke mean score is expected to increase by .313 units, holding the remaining predictors constant (Identity B_1 =.313); for every one-unit increase in the teacher immunity mean score, the teacher stroke mean score is expected to increase by .161 units, holding the remaining predictors constant (Immunity B_2 =.161); for every one unit increase on the TPACK mean score, the teacher stroke mean score is expected to increase by .130 units, holding the remaining predictors constant (TPACK B_3 =.130).

The regression results revealed that the model accounted for 46% of the variance and served as a significant predictor of Iranian EFL participants' teacher stroke, $F(3, 120) = 34.552$, $P < .000$. In the analysis, teacher identity, immunity, and TPACK contributed significantly to the model (B_1 =.313; B_2 =.161; B_3 =.130) respectively.

5. Discussion

The present study, as its first aim, investigated the associations among Iranian EFL teachers' stroke, identity, immunity, and TPACK. The findings provided robust statistical evidence indicating significant and positive correlations among these key

variables. Specifically, teachers' stroke, conceptualized as appreciating and recognizing students' presence, was strongly correlated with identity, TPACK, and immunity, while identity, TPACK, and immunity also exhibited moderate to strong intercorrelations. These results suggest that teachers who actively acknowledge and appreciate their students tend to develop a stronger professional identity, possess higher TPACK knowledge, and demonstrate greater immunity.

The findings align with prior research emphasizing the interconnected nature of teacher stroke, professional identity, technological pedagogical expertise, and psychological resilience. The strong correlation between teachers' stroke and immunity corroborates with that of Mercer and Gregersen (2020) who emphasize that teachers' emotional engagement with students significantly contributes to their ability to handle professional challenges effectively. The observed relationship between teachers' stroke and TPACK also aligns with prior research suggesting that when teachers appreciate their students' value and progress, they are more likely to explore and integrate innovative teaching methodologies (Khorsand & Modarresi, 2023; McCroskey, 1998). This appreciation fosters a sense of purpose, motivating them to refine their pedagogical and technological competencies to enhance student engagement.

Similarly, the correlation between teachers' professional identity and TPACK supports findings from Derakhshan et al. (2021), which suggest that teachers who have a well-established professional identity are more confident in integrating technology into their instruction. Teachers with a strong sense of identity perceive technological advancements as tools to enhance student learning rather than as barriers, making them more willing to incorporate innovative educational technologies. Moreover, the association between teachers' professional identity and psychological immunity is in line with Hiver and Dörnyei's (2017) research, which posits that teachers with a strong professional self-concept are better equipped to manage stress and remain resilient in the face of professional adversity. Given the emotional demands of teaching, particularly in the Iranian EFL context, a well-established identity acts as a protective factor that fosters teachers' psychological endurance.

As its second aim, the current research revealed that the teacher identity, teacher immunity, and teacher TPACK as predictors, could all significantly predict the teacher stroke, as the dependent variable of the study. The concept of teacher stroke as explored in association with the teacher identity has not been investigated in the previous literature, according to the researcher's best knowledge. Nonetheless, the relationship between the teacher's identity and the teacher's stroke can be elucidated by the theoretical framework of Transactional Analysis (TA) in the context of interpersonal relationships in education (Berne, 1988). In this respect, TA may be characterized as "a theory of personality and a structured psychotherapy for personal transformation and progress" (Stewart & Joines, 1987, p. 3). According to the theory, a "stroke" is any action a person does to demonstrate that they are aware of the people around them (Shirai, 2006). In conclusion, the behaviors of giving and getting strokes may satiate a person's insatiable emotional need for acknowledgment and praise from others (Berne, 1988). Accordingly, It could be argued that the stroking behaviors of EFL teachers can contribute to the acknowledgment and recognition of their identity. That is why the teacher's identity as a concept in which the teacher's stroking behavior could be embedded can predict the teacher's stroking behavior.

The predictive relationship between EFL teacher TPACK and their stroke is also an unexplored research line as the review of the past literature indicated. The predictive relationship, however, could be accounted for through the teacher credibility theory (McCroskey & Teven, 1999). The degree to which a teacher is seen as trustworthy, dependable, or credible when applied to the educational setting, is referred to as teacher credibility (McCroskey, 1998). According to McCroskey and Teven (1999), the three components of teacher credibility, namely goodwill, competence, and trustworthiness respectively refer to instructors' concern for students' best interests, expertise in their field, and honesty. Accordingly, it could be argued that for an EFL teacher to have successful stroking behaviors, the teacher needs to have good knowledge of the subject matter, including TPACK, to serve best in an educational context. The teacher's TPACK can function as a tool in her hands to fulfill students' educational needs and consequently gain greater credibility in the educational context. This credibility can pave the way for the teacher to provide more stroke that in a reciprocal relationship can result in more credibility. Therefore, a teacher's TPACK can act as a predictor of the stroking behaviors an EFL teacher practices in instructional settings.

The past literature on the association between the instructor's immunity and the instructor's stroke is in its infancy yielding no previous study in this regard. However, the novel predictive association between the two concepts found in the current study can be consistent with the statement that "teacher immunity affects almost everything that teachers do in their careers" (Hiver, 2015, p. 226). Moreover, the predictive/associative relationship between EFL teachers' immunity and their stroke can be explained by the mediating role that the teacher's stroke can play with other variables that are correlated with the teacher's immunity. For example, teacher stoke as a mediating variable can explain the relationship between EFL instructors' engagement and their immunity. In this respect, Noughabi et al.'s (2020) findings demonstrated that "various aspects of teacher engagement (i.e., cognitive engagement, emotional engagement, social engagement with learners, and social engagement with colleagues) could affect EFL teachers' immunity" (p. 7). It might be argued that through stroking behaviors, teachers strengthen their engagements with both colleagues and learners, thereby creating a pathway toward developing immunity. However, this mediating role explained for the teacher's stroke needs to be supported by further empirical studies.

6. Conclusion

Instructors, including EFL teachers, have been recognized as significant stakeholders within the educational system, with the ability to influence students' academic advancements and the overall effectiveness of educational systems. The conclusion that instructors' behavioral, psychological, and instructional traits need proper consideration and investigation is the result of their significant influence on a range of student-related academic choices (Burroughs et al., 2019; Derakhshan et al., 2020). As a result,

the present study was an attempt to shed more light on the behavioral, psychological, and instructional qualities of EFL instructors. The current study in its approach to the four main variables of the research was novel in Iran.

The findings of the present study provide strong empirical support for the significant relationships among Iranian EFL teachers' stroke, identity, immunity, and TPACK. The study highlights that teachers who actively recognize and appreciate their students' presence tend to develop a more robust professional identity, possess higher technological pedagogical content knowledge, and demonstrate greater psychological resilience. Moreover, teacher identity, immunity, and TPACK were found to significantly predict teacher stroke, underscoring the interconnected nature of these variables. These results contribute to the growing body of research on teacher psychology and professional development, offering new insights into how teacher acknowledgment of students can enhance instructional effectiveness and teacher well-being.

7. Implications

Considering the results found in the present research, this kind of study has important implications for the practice of pre- and in-service EFL instructors, instructor trainers, and administrators in charge of instructor employment initiatives, among other stakeholders in EFL settings. It is obvious that teachers are important to education and that both internal and external variables affect their performance. Hence, students' motivation, learning, achievements, and the educational system efficacy in general, all depend on several teacher's attributes, including but not limited to the teacher's stroke, teacher's immunity, teacher's TPACK, and teacher's Identity, which are shaped by behavioral, psychological, and instructional factors (Derakhshan et al., 2021). EFL instructors must be trained and mentally equipped to handle challenges since teaching is a difficult and stressful profession that puts strain on them. By learning about the demands of teaching languages and the fact that this line of work is well known for its emotional strains, adversities, and stress, in addition to the content, pedagogical, and technological knowledge challenges that EFL teachers may face, teachers can benefit from this line of research. Similar to this, EFL instructors may build an affluent environment for their educational settings where EFL teachers can have professional growth and EFL learners can experience excellent education by learning how to cope with these challenges and demands via developing necessary techniques.

Another implication of the study is to recommend teacher educators provide content and pedagogical knowledge in their teacher education programs with pre-service and in-service teachers regarding the concepts of the teacher's stroke, teacher's identity, teacher's immunity, teacher's TPACK as they are correlated with one another, according to the results obtained in the current study. They can learn the concepts, negotiate different aspects of the concepts, and develop their basic knowledge for their future careers in teacher education programs. The presentation of such concepts and how important they could be in teacher education programs can prepare the teachers to enact their professional knowledge in their instructional settings, paving the way to have more successful English language in EFL contexts.

The findings of this research can also be advantageous for individuals responsible for hiring qualified EFL instructors for educational programs. Only EFL teachers who exhibit a professional identity with stable immunity and those who possess a sufficient understanding of various facets of teacher competence, such as knowledge of stroke and TPACK, among other content, language, psychology, sociology, instructional communication, and cultural knowledge, are allowed to pass the gate in this regard.

8. Limitations and suggestions for further research

Despite its valuable contributions, this study has certain limitations that should be acknowledged. First, the study focused exclusively on Iranian EFL teachers, which may limit the generalizability of the findings to different cultural or educational contexts. Future research could extend this investigation to diverse linguistic and geographical settings to validate the results further. Second, the study employed a quantitative design, which, while effective in establishing statistical relationships, does not capture the in-depth qualitative experiences of teachers regarding stroke, identity, and immunity. Future research could employ qualitative methodologies, such as interviews or classroom observations, to gain deeper insights into these dynamics. Additionally, the study relied on self-reported data, which may introduce response bias. Future studies could use a mixed-methods approach to triangulate findings and enhance validity. Finally, experimental studies examining the effects of targeted teacher training interventions on increasing teacher stroke and its subsequent impact on student engagement and achievement would offer practical applications for teacher education and professional development programs.

9. References

Afsharpour, R., Barani, G., Seyyedrezaei, S. H., & Seyyedrezaei, Z. S. (2023). Exploring the Interplay Among Stroke, Strictness, and Teacher Success from EFL Teachers' Perspective: A Mixed-Methods Study. *Language and Translation, 13*(4), 221-239. https://doi.org/10.30495/ttlt.2024.706672

Aneja, G. A. (2016). (Non) native speakers: Rethinking (non) nativeness and teacher identity in TESOL teacher education. *Tesol Quarterly, 50*(3), 572-596. http://dx.doi.org/10.1002/tesq.315

Atai, M. R., Nazari, M., & Hamidi, F. (2022). Novice EAP teacher identity construction: A qualitative study from Iran. *Journal of English for Academic Purposes, 59*, 101162. http://dx.doi.org/10.1016/j.jeap.2022.101162

Azari Noughabi, M., Ghonsooly, B., & Jahedizadeh, S. (2024). Modeling the associations between EFL teachers' immunity, L2 grit, and work engagement. *Journal of Multilingual and Multicultural Development*, *45*(8), 3158-3173. http://dx.doi.org/10.1080/01434632.2022.2088766

Bagheri, M. (2020). Validation of Iranian EFL Teachers' Technological Pedagogical Content Knowledge (TPACK) Scale. *TESL-EJ*, *24*(2), 1-20.

Berne, E. (1988). *Games people play: The psychology of human relationships*. Grove Press.

Bostancıoğlu, A., & Handley, Z. (2018). Developing and validating a questionnaire for evaluating the EFL 'Total PACKage': Technological Pedagogical Content Knowledge (TPACK) for English as a Foreign Language (EFL). *Computer Assisted Language Learning*, *31*(5–6), 572–598. https://doi.org/10.1080/09588221.2017.1422524

Bressler, C., & Rotter, C. (2017). The Relevance of a Migration Background to the Professional Identity of Teachers. *International Journal of Higher Education*, *6*(1), 239-250.

Burroughs, N., Gardner, J., Lee, Y., Guo, S., Touitou, I., Jansen, K., & Schmidt, W. (2019). *Teaching for excellence and equity: Analyzing teacher characteristics, behaviors and student outcomes with TIMSS*. Springer nature.

Chai, C. S., Chin, C. K., Koh, J. H. L., & Tan, C. L. (2013). Exploring Singaporean Chinese language teachers' technological pedagogical content knowledge and its relationship to the teachers' pedagogical beliefs. *The Asia-Pacific Education Researcher*, *22*, 657-666. https://doi.org/10.1007/s40299-013-0071-3

Cohen, J. (1988). *The effect size. Statistical power analysis for the behavioral sciences.* Routledge.

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

De Costa, P. I., & Norton, B. (2017). Introduction: Identity, transdisciplinarity, and the good language teacher. *The Modern Language Journal*, *101*(S1), 3-14. https://doi.org/10.1111/modl.12368

Derakhshan, A., Coombe, C., Zhaleh, K., & Tabatabaeian, M. (2020). Examining the Roles of Continuing Professional Development Needs and Views of Research in English Language Teachers' Success. *Tesl-Ej*, *24*(3), 1-27.

Derakhshan, A., Doliński, D., Zhaleh, K., Janebi Enayat, M., and Fathi, J. (2021). *A cross-cultural study of Iranian and Polish higher education students' academic engagement in terms of teacher care and teacher-student rapport*. Gorgan: Golestan University, Grant Number: 992151.

Freedman, A. (1993). Show and tell? The role of explicit teaching in the learning of new genres. *Research in the Teaching of English*, 222-251. https://doi.org/10.58680/rte199315402

Frymier, A. B., Goldman, Z. W., & Claus, C. J. (2019). Why nonverbal immediacy matters: A motivation explanation. *Communication Quarterly*, *67*(5), 526-539. https://doi.org/10.1080/01463373.2019.1668442

Ghiasvand, F., Kogani, M., & Nemati, F. (2023). "Teachers as conflict managers": mapping novice and experienced Iranian EFL teachers' professional identity conflicts and confrontation strategies. *Asian-Pacific Journal of Second and Foreign Language Education*, *8*(1), 45-64. https://doi.org/10.1186/s40862-023-00219-z

Gooran, M., Soleimani, H., Alavi, M., & Jafarigohar, M. (2023). EFL teachers' immunity: a case of online language teaching. *Journal of Multilingual and Multicultural Development*, *44*(10), 908-927. https://doi.org/10.1080/01434632.2022.2044341

Hahs-Vaughn, D. L., & Lomax, R. G. (2020). Multiple linear regression. *Statistical Concepts*, 527-599.

Haseli Songhori, M., Ghonsooly, B., & Afraz, S. (2018). Language teacher immunity among Iranian EFL teachers: A self-organization perspective. *Iranian Journal of English for Academic Purposes*, *7*(1), 128-143.

Hayes, A. F., & Cai, L. (2007). Using heteroskedasticity-consistent standard error estimators in OLS regression: An introduction and software implementation. *Behavior research methods*, *39*, 709-722. https://doi.org/10.3758/BF03192961

Hiver, P. (2015). Once burned, twice shy: The dynamic development of system immunity in language teachers. In Z. Dörnyei, P. D. MacIntyre, & A. Henry (Eds.), Motivational Dynamics in Language Learning. Multilingual Matters (pp. 214-237). https://doi.org/10.21832/9781783092574-017

Hiver, P. (2016). *Tracing the signature dynamics of language teacher immunity* Unpublished Ph.D. Thesis, University of Nottingham, England.

Hiver, P. (2017). Tracing the signature dynamics of language teacher immunity: A retrodictive qualitative modeling study. *The Modern Language Journal*, *101*(4), 669-690. https://doi.org/10.1111/modl.12433

Hiver, P., & Dörnyei, Z. (2017). Language teacher immunity: A double-edged sword. *Applied Linguistics*, *38*(3), 405-423. https://doi.org/10.1093/applin/amv034

Huang, I. C. (2014). Contextualizing teacher identity of non-native-English speakers in US secondary ESL classrooms: A Bakhtinian perspective. *Linguistics and Education*, *25*, 119-128. https://doi.org/10.1016/j.linged.2013.09.011

Irajzad, F., & Shahriari, H. (2017). Student socioeconomic status and teacher stroke: A case of female students in Iran. *Discourse and Communication for Sustainable Education*, *8*(1), 129-144. http://dx.doi.org/10.1515/dcse-2017-0010

Jiang, L., Yuan, K., & Yu, S. (2021). Transitioning from pre-service to novice: A study on Macau EFL teachers' identity change. *The Asia-Pacific Education Researcher*, *30*, 11-21. https://doi.org/10.1007/s40299-020-00510-4

Johnston, B., Pawan, F., & Mahan-Taylor, R. (2005). The professional development of working ESL/EFL teachers: A pilot study. *Second language teacher education: International perspectives*, 53-72.

Kadkhodaie, Sh., Talebinejad, M. R., & Shahrokhi, M. (2023). *Iranian EFL Teachers' Immunity, Technological Pedagogical Content Knowledge (TPACK), Identity Construction, and Stroke.* Unpublished Ph. D. Thesis, Shahreza Branch, Islamic Azad Univerity, Iran

Kalali Sani, S. F., Motallebzadeh, K., Khodabakhshzadeh, H., & Zeraatpisheh, M. (2021). Iranian EFL teachers' professional identity and their goal orientation. *Teaching English Language*, *15*(1), 137-160. https://doi.org/10.22132/tel.2021.132248

Kavanoz, S., Yüksel, H. G., & Özcan, E. (2015). Pre-service teachers' self-efficacy perceptions on Web Pedagogical Content Knowledge. *Computers & Education*, *85*, 94-101. https://doi.org/10.1016/j.compedu.2015.02.005

Kayi-Aydar, H. (2015). Multiple identities, negotiations, and agency across time and space: A narrative inquiry of a foreign language teacher candidate. *Critical Inquiry in Language Studies*, *12*(2), 137-160. https://doi.org/10.1080/15427587.2015.1032076

Khorsand, M., & Modarresi, G. (2023). The relationship between teachers' emotions, strokes, and academic achievement: The case of BA English-major students. *Language and Translation Studies*, *56*(2), 71-107. https://doi.org/10.22067/lts.2023.81620.1179

Koehler, M., & Mishra, P. (2009). What is technological pedagogical content knowledge (TPACK)? *Contemporary issues in technology and teacher education*, *9*(1), 60-70.

Kyriacou, C. (2001). Teacher stress: Directions for future research. *Educational Review*, *53*(1), 27-35. https://doi.org/10.1080/00131910120033628

Liou, I. Y. S. (2008). *English as an international language and teachers' professional identity.* Unpublished Ph. D. Thesis, Deakin University, England.

Liu, M. H., & Kleinsasser, R. C. (2015). Exploring EFL teachers' CALL knowledge and competencies: In-service program perspectives. *Language Learning & Technology*, *19*(1), 119–138.

Low, M., Davey, H., & Davey, J. (2012). Tracking the professional identity changes of an accountancy institute: The New Zealand experience. *Journal of Accounting & Organizational Change*, *8*(1), 4-40. https://doi.org/10.1108/18325911211205720

McCroskey, J. C. (1998). *An introduction to communication in the classroom* (2nd ed.). Tapestry Press.

McCroskey, J. C., & Teven, J. J. (1999). Goodwill: A reexamination of the construct and its measurement. *Communications Monographs*, *66*(1), 90-103. https://doi.org/10.1080/03637759909376464

Mei, B., Brown, G. T., & Teo, T. (2018). Toward an understanding of preservice English as a foreign language teachers' acceptance of computer-assisted language learning 2.0 in the People's Republic of China. *Journal of Educational Computing Research*, *56*(1), 74-104. https://doi.org/10.1177/0735633117700144

Mercer, S. (2018). Psychology for language learning: Spare a thought for the teacher. *Language Teaching*, *51*(4), 504-525. https://doi.org/10.1017/S0261444817000258

Mercer, S., & Gregersen, T. (2020). *Teacher wellbeing*. Oxford University Press.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge. *Teachers College Record*, *108*(6), 1017-1054. https://doi.org/10.1111/j.1467-9620.2006.00684.x

Moradkhani, S., & Ebadijalal, M. (2024). Professional identity development of Iranian EFL teachers: Workplace conflicts and identity fluctuations. *Journal of Language, Identity & Education*, *23*(2), 256-270. https://doi.org/10.1080/15348458.2021.1988605

Morgan, B., & Clarke, M. (2011). *Identity in second language teaching and learning.* Routledge.

Najjari, R., Abbasian, G. R., & Yazdanimoghaddam, M. (2021). Assessment and development of Iranian EFL teachers' technological pedagogical content knowledge (TPACK). *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, *40*(4), 161-193. https://doi.org/10.22099/jtls.2021.39038.2914

Namaghi, F. R. (2016). *The influence of stroke on "willingness to attend classes"(WTAC) and foreign language achievement: A case of Iranian EFL learners.* Unpublished Master's Thesis, Ferdowsi University of Mashhad, Iran.

Namaziandost, E., Heydarnejad, T., & Rezai, A. (2023). Iranian EFL teachers' reflective teaching, emotion regulation, and immunity: examining possible relationships. *Current Psychology*, *42*(3), 2294-2309. https://doi.org/10.1007/s12144-022-03786-5

Nessaei, H., & Kazemi, S. (2024). On the Relationship between Iranian EFL Teachers' Technological Pedagogical Content Knowledge (TPCK) and their Willingness to Attend Continuing Professional Development (CPD) Courses. *Teaching and Learning*, *4*(1). 27-45.

Noorbakhsh, Z., Pishghadam, R., & Saboori, F. (2018). Stroke and gender identity in teacher success: From learners' viewpoints. *Sri Lanka Journal of Social Sciences*, *41*(1), 39-48.

Noughabi, M. A., Amirian, S. M. R., Adel, S. M. R., & Zareian, G. (2020). The association of experienced in-service EFL teachers' immunity with engagement, emotions, and autonomy. *Current Psychology*, *1*-10. https://doi.org/10.1007/s12144-020-01066-8

Park, P. H. I. L. I. P. P. A. (2017). Understanding the connections between second language teacher identity, efficacy, and attrition: a critical review of recent literature. *Journal of Belonging, Identity, Language, and Diversity*, *1*(1), 75-91.

Peng, J. E., & Woodrow, L. (2010). Willingness to communicate in English: A model in the Chinese EFL classroom context. *Language learning*, *60*(4), 834-876. https://doi.org/10.1111/j.1467-9922.2010.00576.x

Pierson, W. J. (2003). *Reflection on the process of learning the work of nurses during practice experiences* Unpublished Ph. D. Thesis, Simon Fraser University, Canada.

Pishghadam, R., & Farkhondefal, E. (2017). From navazesh (physical strokes) to navazeh (mental strokes): A look into the concept of stroking in teaching a second language. *Language and Translation Studies (LTS)*, *49*(4), 1-13. https://doi.org/10.22067/lts.v49i4.62725

Pishghadam, R., & Khajavy, G. H. (2014). Development and validation of the student stroke scale and examining its relation with academic motivation. *Studies in Educational Evaluation*, *43*, 109-114. https://doi.org/10.1016/j.stueduc.2014.03.004

Pishghadam, R., Derakhshan, A., & Zhaleh, K. (2019). The interplay of teacher success, credibility, and stroke with respect to EFL students' willingness to attend classes. *Polish Psychological Bulletin*, *50*(4), 284-292. https://doi.org/10.24425/ppb.2019.131001

Pishghadam, R., Derakhshan, A., Zhaleh, K., & Al-Obaydi, L. H. (2023). Students' willingness to attend EFL classes with respect to teachers' credibility, stroke, and success: a cross-cultural study of Iranian and Iraqi students' perceptions. *Current Psychology, 42*(5), 4065-4079. https://doi.org/10.1007/s12144-021-01738-z

Pishghadam, R., Derakhshan, A., Zhaleh, K., & Al-Obaydi, L. H. (2021). Students' willingness to attend EFL classes with respect to teachers' credibility, stroke, and success: A cross-cultural study of Iranian and Iraqi students' perceptions. *Current Psychology*. https://doi.org/10.1007/s12144-021-01738-z

Pishghadam, R., Ghonsooly, B., & Zhaleh, K. (2018). Effects of conceptions of intelligence and ambiguity tolerance on teacher burnout: A case of Iranian EFL teachers. *Journal of Research in Applied Linguistics*, *9*(2), 118-140. https://doi.org/10.22055/rals.2018.13796

Pituch, K. A., & Stevens, J. P. (2016). *Applied Multivariate Statistics for the Social Sciences.* 6th ed. New York and London.

Rahmati, T., Sadeghi, K., & Ghaderi, F. (2019). English as a foreign language teacher immunity: An integrated reflective practice. *Iranian Journal of Language Teaching Research*, *7*(3 (Special Issue)), 91-107. https://doi.org/10.30466/ijltr.2019.120738

Rajabnejad, F., Pishghadam, R., & Saboori, F. (2017). On the influence of stroke on willingness to attend classes and foreign language achievement. *Applied Research on English Language*, *6*(2), 141-158. https://doi.org/10.22108/are.2017.21344

Raygan, A., & Moradkhani, S. (2022). Factors influencing technology integration in an EFL context: investigating EFL teachers' attitudes, TPACK level, and educational climate. *Computer Assisted Language Learning*, *35*(8), 1789-1810. https://doi.org/10.1080/09588221.2020.1839106

Rodgers, C. R., & Scott, K. H. (2008). The development of the personal self and professional identity in learning to teach. In *Handbook of research on teacher education* (pp. 732-755). Routledge.

Rudolph, N. (2016). Negotiating borders of being and becoming in and beyond the English language teaching classroom: Two university student narratives from Japan. *Asian Englishes*, *18*(1), 2-18. https://doi.org/10.1080/13488678.2015.1132110

Saudelli, M. G., & Ciampa, K. (2016). Exploring the role of TPACK and teacher self-efficacy: An ethnographic case study of three iPad language arts classes. *Technology, Pedagogy, and Education*, *25*(2), 227-247. https://doi.org/10.1080/1475939X.2014.979865

Sauter D. A. (2017). The Nonverbal Communication of Positive Emotions: An Emotion Family Approach. *Emotion review: Journal of the International Society for Research on Emotion*, *9*(3), 222–234. https://doi.org/10.1177/1754073916667236

Sedikides, C., & Brewer, M. B. (2015). *Individual self, relational self, and collective self*. Psychology Press.

Shirai, S. (2006) How transactional analysis can be used in terminal care. *International Congress Series*, 1278, 179-184. DOI: https://doi.org/10.1016/j.ics.2005.12.052

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. https://doi.org/10.3102/0013189X015002004

Skorikov, V. B., & Vondracek, F. W. (2011). Occupational identity. *Handbook of identity theory and research*, 693-714.

Song, W., Zhao, T., Huang, E., & Liu, W. (2022). How Positive and Negative Emotions Promote Ritualistic Consumption Through Different Mechanisms. *Frontiers in Psychology*, *13*, 901572. https://doi.org/10.3389/fpsyg.2022.901572

Stewart, I., & Joines, V. (1987). *TA today: A new introduction to transactional analysis*. Nottingham: Lifespace Pub.

Trent, J. (2017). Discourse, agency and teacher attrition: Exploring stories to leave by amongst former early career English language teachers in Hong Kong. *Research Papers in Education*, *32*(1), 84-105. https://doi.org/10.1080/02671522.2016.1144215

Tsakissiris, J. (2015). *The role of professional identity and self-interest in career choices in the emerging ICT workforce*. Unpublished Master's. Thesis, The Queensland University of Technology. Australia.

Tseng, J. J., Chai, C. S., Tan, L., & Park, M. (2020). A critical review of research on technological pedagogical and content knowledge (TPACK) in language teaching. *Computer Assisted Language Learning*, *35*(4), 948-971. https://doi.org/10.1080/09588221.2020.1868531

Tsui, A. (2003). *Understanding expertise in teaching: Case studies of second language teachers*. Cambridge University Press.

Vitanova, G. (2016). Exploring second-language teachers' identities through multimodal narratives: Gender and race discourses. *Critical Inquiry in Language Studies*, *13*(4), 261-288. https://doi.org/10.1080/15427587.2016.1165074

Wang, Y., Derakhshan, A., & Azari Noughabi, M. (2024). The interplay of EFL teachers' immunity, work engagement, and psychological well-being: Evidence from four Asian countries. *Journal of Multilingual and Multicultural Development*, *45*(8), 3241-3257. https://doi.org/10.1080/01434632.2022.2092625

Wolff, D., & De Costa, P. I. (2017). Expanding the language teacher identity landscape: An investigation of the emotions and strategies of an NNEST. *The Modern Language Journal*, *101*(S1), 76-90. https://doi.org/10.1111/modl.12370

Yang, S. (2018, January). Riding the waves of augmented reality in preservice EFL teacher education: Model development. In *Proceedings of the 6th International Conference on Information and Education Technology* (pp. 185-189). https://doi.org/10.1145/3178158.3178168

Yazan, B. (2017). "It just made me look at language in a different way:" ESOL teacher candidates' identity negotiation through teacher education coursework. *Linguistics and Education*, *40*, 38-49. https://doi.org/https://doi.org/10.1016/j.linged.2017.06.002

Yazan, B. (2018). TESL Teacher Educators' Professional Self-Development, Identity, and Agency. *TESL Canada Journal*, *35*(2), 140–155. https://doi.org/10.18806/tesl.v35i2.1294

Yazan, B. (2023). A conceptual framework to understand language teacher identities. *Second Language Teacher Education*, *1*(2), 185-208. https://doi.org/10.1558/slte.24908

Yazdanpour, H. (2015). *Constructing and validating a teacher stroke scale and examining its relationship with burnout.* Unpublished Master's Thesis, Ferdowsi University of Mashhad, Mashhad, Iran.

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 147-161

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.8592.1058

Scaffolding Through Google Docs and Writing Achievement: The Mediating Role of Metacognitive Knowledge

Mojgan Amiri[1]
Majid Farahian[2*]

[1]M.A. Student of Applied Linguistics, Department of ELT, Ker.C., Islamic Azad University, Kermanshah, Iran
[*2]Assistant Professor of Applied Linguistics, Department of ELT, Ker.C., Islamic Azad University, Kermanshah, Iran

ARTICLE HISTORY

Received: 16 December 2024
Revised: 19 January 2025
Accepted: 11 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: majid.farahian@iau.ac.ir

ABSTRACT

Previous studies have explored the significant function of writing in the language development process, the essential role of metacognition in learning, the efficacy of scaffolding in developing students' writing skill, and the application of technology in the learning and training process. However, the relationship between writing metacognitive knowledge and scaffolding via Google Docs and academic writing achievement needs further investigation. Therefore, the present study sought to examine whether writing metacognitive knowledge mediates the relationship between English as a foreign language (EFL) learners' scaffolding via Google Docs and academic writing achievement. The participants were 90 EFL undergraduate students enrolled in a writing course. The proficiency level of the participants was assessed through an online DIALANG test. Then, a quantitative correlational design was used to test the research hypothesis. Two questionnaires were used to gather quantitative data, and students' performance on academic writing achievement was assessed. To evaluate the data, Pearson correlation, multiple linear regression, simple linear regression, and path analysis were run. The findings uncovered the hypothesized model of relationships among the variables of the study. More specifically, the results supported the mediator role of writing metacognitive knowledge.

KEYWORDS: EFL writing skill; Peer feedback; Technology-mediated writing; Writing metacognitive knowledge

1. Introduction

Writing is a fundamental communication skill for learning a second language and is critical for language learners' achievement (Geng et al., 2021). It is more than a way of communicating in a foreign language because it can help learners acquire the foreign language (Maftoon et al., 2014). Moreover, the act of writing is a creative process through which one creates a text for another person to read (Nosratinia & Adibifar, 2014; Zarinkamar et al., 2021), hence helping them to communicate. In the writing process, there are stages of constructing the text, memory and executive functions (Owens, 2012) and it has been assumed that metacognition as an executive function regulates and controls the process of writing (Ramadhanti & Yanda, 2021). It has been claimed that successful writing as a very complex cognitive process that involves planning, text generation, and revision (Flower et al., 1994) depends on metacognitive knowledge and metacognitive regulation behaviour (McCormick, 2003).

Another variable discussed in the present study is the use of Google Docs in EFL writing courses. A growing body of research supports the incorporation of Web 2.0 applications such as forums, blogs, wikis, and Google Docs in L2 writing courses, as these tools enable educators to establish interactive online environments for collaborative writing, where L2 writers can receive feedback from both instructors and peers, as well as participate in peer feedback and peer editing processes (Dizon, 2016; Strobl, 2013). Nevertheless, the exploration of Google Docs in writing courses has been less extensively researched compared to other Web 2.0 tools (i.e., blogs or wikis) (Ebadi & Rahimi, 2017). Considering that Google Docs presents several useful features for writing courses, such as peer feedback, peer editing, redrafting, and text change tracking (Semeraro & Moore, 2016), gathering more

Cite this article: Amiri, M. and Farahian, M. (2025). Scaffolding through Google Docs and writing achievement: The mediating role of metacognitive knowledge. *Applied Linguistics Inquiry*, *3*(1), 147-161. doi: 10.22077/ali.2025.8592.1058

empirical evidence regarding the effectiveness of Google Docs could offer a promising and practical technological application for EFL writing instruction.

Metacognitive abilities are essential for EFL learners' academic achievements (Cai & Zhao, 2023; Sun et al., 2024). They empower students to go beyond being mere recipients of information, allowing them to engage critically with the content, establish achievable goals, and modify their learning strategies for greater effectiveness. In the current digital era, marked by an overflow of problems and frequently conflicting information sources, metacognitive skills have become increasingly vital for evaluating credible information and for pursuing self-directed, lifelong learning. Research has also indicated the role of Artificial Intelligence (AI) in improving metacognitive, creative skills and higher order thinking (Ilgun Dibek et al., 2024; Khotimah, et al., 2024). In a meta-analysis, Ilgun Dibek et al. (2024) reported that the use of AIEd had a positive impact on higher order thinking skills. In the same vein, Darwin et al. (2023) fond that Artificial intelligence can be beneficial in fostering critical thinking abilities, but, as they report, there are caveats that need to be managed carefully. Szmyd and Mitera (2024) also hold that learners recognize the significance of critically evaluating their own beliefs as well as those of others. They emphasize that while AI can assist in this evaluation, it cannot substitute for conventional teaching approaches, which are crucial for developing autonomous thinking.

Although a lot of researchers have acknowledged the key role of metacognition in facilitating writing in the EFL context (Farahian & Avarzamani, 2018; Sun & Zhang, 2022; Wang et al., 2024), there is limited information regarding the interface among scaffolding through Google Docs, metacognitive knowledge, and EFL writing achievement. Therefore, the present study, as its main objective, sought to investigate the mediating role of writing metacognitive knowledge in the relationship of scaffolding via Google Docs and academic writing achievement.

The present study is significant since it attempts to bridge the gap between technology-prosperous teaching and metacognitive development in EFL writing. By investigating how scaffolding via Google Docs impacts metacognitive knowledge and writing achievement, this research offers insights into optimizing digital tools for cognitive and metacognitive growth. In the era of increasing AI and digital integration in education, understanding this interplay is important to design effective instructional strategies. This study contributes to both theoretical knowledge and practical applications, which emphasizes the importance of promoting self-regulations and important thinking skills required to learn academic success and lifelong learning in the digital age.

2. Literature review

2.1. Metacognition

According to Flavell (1987), metacognitive knowledge is "the part of one's acquired word knowledge that has to do with cognitive matters" (p. 21). Generally speaking, metacognition is considered as one's thinking about his/her thinking (Dennis & Somerville, 2022) and is defined as the knowledge about and regulation of cognitive processes (Yu-Ling et al., 2001). When learners get involved in planning, monitoring, evaluating, and making changes to their own learning behaviours, metacognition comes into play (Vakilifard & Abedini, 2021). As Flavell (1979) suggests, metacognition is knowing about knowing or cognition about cognition.

Metacognition and its function in the development of writing have been investigated by various researchers (e.g., Colognesi et al., 2020; Ramadhanti & Yanda, 2021). Furthermore, they have scrutinized the correlation between metacognitive skills and writing achievement (e.g., Eriyani, 2020; Teng, 2019). Therefore, metacognition-affected training is distinguished as an efficient teaching procedure (Al-Jarrah et al., 2018).

Even though students benefit from developing this knowledge, particularly in writing courses, they are not taught how to cultivate metacognitive skills, especially in FL contexts like Iran (Nourazar et al., 2022) where product-oriented oriented approaches to learning and teaching are valued (Cheraghi et al., 2022). In this context many teachers assume that the instruction of grammar and vocabulary would be sufficient for writing and therefore the significant role of writing processes has been disregarded in EFL classes (Avarzamani & Farahian, 2019).

2.2. Technology in education

Nowadays, technology has advanced tremendously, enabling the concept of scaffolding—originally limited to human interactions—to expand beyond human agents since the early 1990s. This evolution now incorporates web-based tools as non-human scaffolding supports in learning environments (Kadkhodaei et. al, 2025). Although the notion of collaborative writing has been around for some time, its devices have undergone significant modifications. Research shows that online cooperative activities (e.g., blogs & wiki, an essential component of Web 2.0) is contingent upon the improvement of three aptitudes among college students: (1) social skills (Apple et al., 2011), (2) teamwork (Blair, 2006), (3) and basic computing skills (Bottge et al., 2009). Collaborative tasks enable students to learn beyond what they might have learned alone, to share perspectives, and to complete tasks successfully, which can maximize learning both inside and outside of the classroom (Jones, 2007) In this regard, Google Docs is a platform for online collaborative writing. It enables numerous authors to edit in real-time and facilitates collaborative writing with the ease of accessing it from anywhere. In this context, Hemati and Farahian (2024) notes that Google Docs as an online digital media helps instructors in collaborative writing by preparing effective features that aid students in developing collaborative writing skills. Google Docs enables people to collaborate on the same task without limitations often created by usual face-to-face interactions (Perron & Sellers, 2011).

Research on the use of digital tools for scaffolding has shown promising results to improve both writing achievement and metacognition. For example, the study on collaborative writing platforms such as Google Docs has demonstrated that these tools can enhance the writing achievement by providing real-time feedback, enabling peer collaboration, and offering structured support (Fathi et al., 2021; Graham & Perin, 2007; Nabhan & Sa'diyah, 2021; Nhung & Hue, 2022; Mahmood, 2018). Similarly, scaffolding through digital tools have been found to promote metacognition by encouraging learners to reflect on their writing processes, set goals, and monitor their progress (Cho & McArther, 2010; Zheng et al., 2015). While specific studies on 'scaffolding through Google Docs' are limited, comprehensive literature on technology-mediated scaffolding shows that such tools can support both writing achievement and metacognitive development by creating environment for interactive and adaptive learning. Despite the growing body of research in this area, only some studies have concentrated on the connections between students' metacognitive awareness and their writing performance (e.g. Teng, 2019) and this has been particularly the case in EFL contexts. Furthermore, no significant research has reported the interface among scaffolding through Google Docs, metacognitive knowledge, and EFL writing achievement. In an effort to address this research lacuna, the present study aimed at exploring the mediating effect of writing metacognitive knowledge in the relationship of scaffolding via Google Docs and academic writing achievement. Accordingly, the hypothesized model was presented (see Figure 1) and subsequent research questions were proposed:

1. Is there a significant relationship between EFL learners' scaffolding via Google Docs and their writing achievement?

2. Is there a significant relationship between EFL learners' scaffolding via Google Docs and their writing metacognitive knowledge?

3. Is there a significant relationship between EFL learners' writing metacognitive knowledge and their writing achievement?

4. Does EFL learners' writing metacognitive knowledge mediate the relationship between scaffolding via Google Docs and their writing achievement?

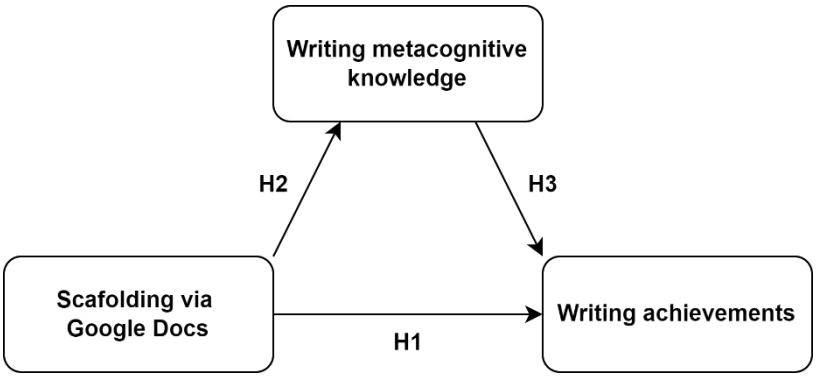

Figure 1. The hypothesized model

3. Methodology

3.1. Design

The current study explored the mediating role of writing metacognitive knowledge in the relationship of scaffolding via Google Docs and academic writing achievement. Therefore, this study utilized a quantitative and correlational design to test research hypotheses.

3.2. Participants

The study involved 90 undergraduate students (56 females and 34 males) who were taking a writing course at a university in Kermanshah (a city in the west of Iran). The ages of the participants varied between 20 and 24 years. To evaluate the participants' proficiency, the online DIALANG test was utilized. The students were categorized into six proficiency levels: A1, A2, B1, B2, C1, and C2, based on their test results. Those students who scored at A1 and A2 levels were excluded from the study since they were unable to handle the tasks involved. Additionally, some participants at the C1 and C2 levels were also removed due to concerns that they may have already developed their metacognitive skills. This exclusion ensured that the study focused on participants who would benefit most from scaffolding via Google Docs and whose metacognitive knowledge was still developing. Finally, from an initial group of 148, 90 students with B1 and B2 English proficiency were selected for the study.

They were in 3 classes with class1 (18 females and 12 males), class2 (17 females and 13 males), and class3 (21 females and 9 males) students, respectively. These classes were chosen because the instructor utilized Google Docs for collaborative writing throughout the term, aligning with the study's focus on scaffolding and metacognitive development. In general terms, the

course aimed to develop students' ability to write 5 paragraph essays. This university was chosen because, following the COVID-19 pandemic, all courses were conducted online, providing a natural setting for examining the impact of scaffolding via Google Docs on writing achievement and metacognitive knowledge.

Before the study, the Scaffolding via Google Docs Questionnaire was administered to inquire whether the students had any writing experience in Google Docs before the study and if they had received any instruction on it. The response we received from all participants was negative. This lack of prior exposure was crucial for isolating the effects of scaffolding via Google Docs on writing achievement and metacognitive knowledge. The participants were also assured that the results would be kept in absolute confidence and would only be applied to the intended research. The first researcher was the instructor and had at least seventeen years of experience instructing EFL in various language schools.

3.3. Instruments

3.3.1. DIALANG Proficiency Test

DIALANG is a diagnostic test. Its main purpose is to inform language learners about their proficiency level. According to the Common European Framework of Reference for Languages, the outcomes of the DIALANG test, which assesses general language proficiency, are classified into levels ranging from A1 to C2.

3.3.2. Writing achievement test

Since it has been suggested that writing is generally simpler when the subject matter is well-known rather than when it is not (McCutchen, 2000) in the final session of the course, students were requested to compose a brief essay ranging from 80 to 120 words on the topic, "Do you believe that social media can have adverse effects on your life?". The scoring rubric developed by Jacobes et al. (1981) was utilized to assess the students' essays. To evaluate students' writing, the rubric emphasizes five areas: organization, content, language use, vocabulary, and mechanics. Four levels are employed to grade each category. Each classification is evaluated using a scale that ranges from very poor to fair, good to average, and excellent to very good. A maximum score of 100 is given to perfect writing.

3.3.3. Metacognitive awareness writing questionnaire

The MAWQ, developed by Farahian (2017), is based on the concept that writing metacognitive knowledge can be divided into two main categories: Knowledge of cognition and regulation of cognition. This tool consists of 36 items and utilizes a 5-point Likert scale, where responses range from strongly agree (1) to strongly disagree (5). The author stated that the scale exhibited a satisfactory level of reliability, with reliability indices ranging from 67 to 91. Regarding the scale's validity, Farahian provides details on the procedure for validating the scale.

3.3.4. Scaffolding via Google Docs questionnaire (SGDQ)

The instrument was designed and validated by Farahian and Ebadi (2022). The framework underlying the development of the questionnaire was informed by the work of Wood et al. (1976), Vygotsky (1986), Hogan and Pressly (1997), Van de Pole et al. (2010), Mortazavi et al. (2016), Ikawati (2020), and Suwastini et al (2021). The SGDQ contains 15 items. It consists of four categories including instructional, interactional, reflective, and affective.

The questionnaire is based on a 5-point Likert scale varying from strongly agree= 5 to strongly disagree= 1. As reported, the questionnaire was examined to estimate the construct validity using the exploratory factor analysis (EFA), and average variance extracted (AVE). In addition, the reliability of the questionnaire was evaluated using Cronbach's alpha coefficient. It was computed to be 0.91.

3.4. Procedure

At the beginning of the study, the DIALANG proficiency test was used to determine the homogeneity of all the students. After the students took the proficiency test, they participated in a writing course. The analysis of the data excluded information regarding students whose proficiency levels fell outside the study's objectives. The first researcher was the teacher of three online classes. During a fourteen-session course, the students were instructed on writing 5 paragraph essays and each session they were given some model paragraphs to analyse during the online courses. They were also given some take-home assignments. The students were instructed to create their own Google Docs and share them with their teammates and the teacher. As part of their assignment, each student was required to choose a partner, read his/her partner's writing task, and provide the necessary feedback. Following Slavin (1996), cooperative learning activities were based on individual accountability, equal participation, simultaneous interaction, and positive interdependence. For example, as he suggests, three conditions in the chosen activities accounted for individual accountability: 1) the student performed individually; 2) group members observed her/his performance, and 3) again the student performed individually. As already explained to the students in two training sessions, when the students performed at the second stage, each pair was monitored performing one or some of the nine language functions for scaffolding

(see, Mohammadzadeh et al., 2022) including agreeing, explaining, giving ideas, instructing, restating, suggesting, comprehension checking, eliciting opinion, and questioning. For example, agreeing, involves expressing agreement with other learners' ideas or explaining the meaning of a term or idea that is not clear to the partner.

The students were expected to synchronously pass comments and edit each other's writing tasks in online classes using Google Docs. To familiarize the students with the process, the first researcher/instructor used some sample films to address the students' possible questions and elaborate on the technical aspects of Google Docs. During the process, each student was asked to revise the writing assignments of another student each time (see Figure 2). The students were required to revise their classmates' work by using a different font colour, focusing on key features like an overview of the content, the accuracy of the information, and the word count (i.e., task achievement); the organization of ideas, paragraph structure, and linking techniques (i.e., coherence and cohesion); and the use of appropriate vocabulary, idiomatic phrases, metaphors, collocations, verbal phrases, prepositions, and other linguistic features. The instructor provided instructional, interactive, reflective, and emotional support to guide students through this process. Each session, the students of three classes received the same number of writing assignments. They were asked to compose a short essay (80–120 words) about a well-known subject that sounded attractive and relevant to the preferences of the students. The students' agreements with the chosen topics were also taken into account. There were 10 consecutive weeks of instruction, each lasting 90 minutes. All of the students took the online exams in the final week of the term. Two instructors who had more experience and were more skilled at teaching composition writing than the researcher/instructor were asked to rate the writing tasks using a scale adapted from Jacobs et al. (1981). This assignment was given to the students in the last session. The findings revealed that the two raters had an inter-rater reliability of .89. Then, students completed two short questionnaires about their views toward Scaffolding via Google Docs and MAWQ at the end of the writing course. Because of the COVID-19 lockdown, the students were not directly accessible to the researchers. Thus, the first researcher delivered the questionnaires via Google Forms.

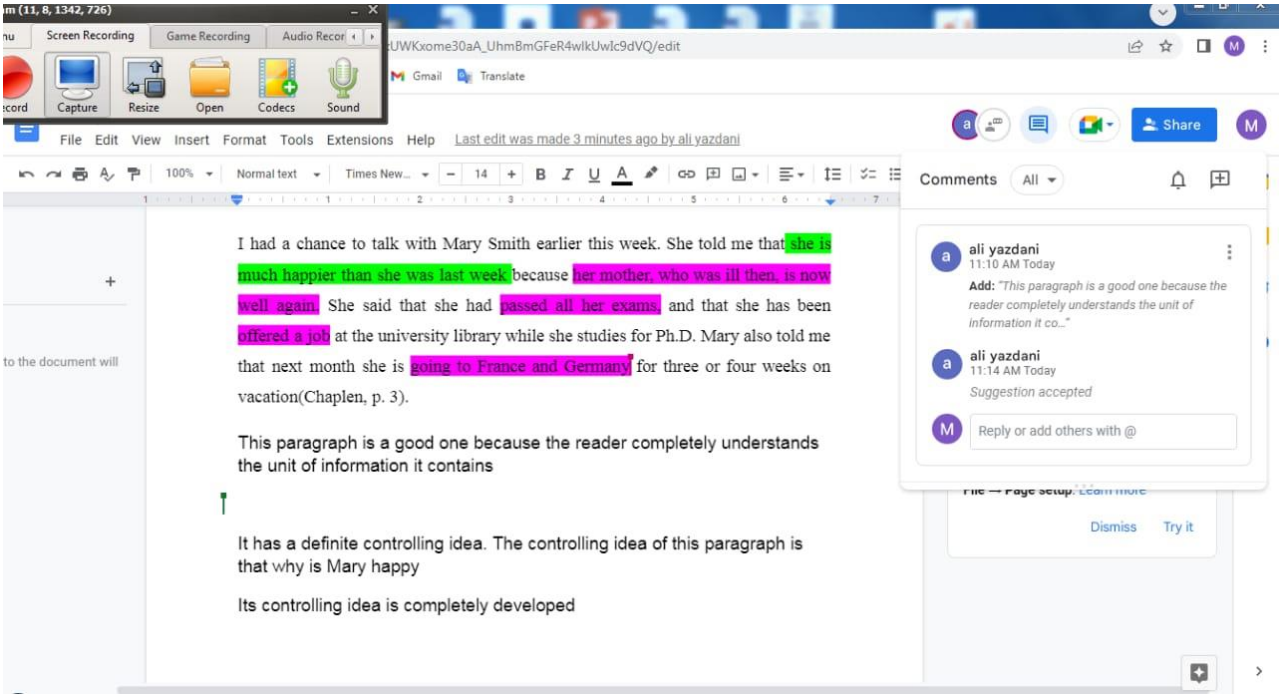

Figure 2. A snapshot of scaffolding via Google Docs

3.5. Data analysis

When scores were allocated to the students' performance on academic writing achievement, and the participants' answers to the items of the questionnaires were analysed quantitatively, using SPSS 23, the researchers ran descriptive statistics to present the statistics related to the variables under the study. In order to further explore the connections between the three variables, three Pearson product-moment correlations were computed. Likewise, to explore the mediating influence of writing metacognitive knowledge on the connections between scaffolding through Google Docs and academic writing achievement, Pearson correlation, multiple linear regression, simple linear regression, and path analysis using Amos version 23 were conducted.

4. Results

Before performing the stages of data analysis, it was necessary to assess the normality of the collected data using One-Sample Kolmogorov-Smirnov Test. The purpose of this test was to demonstrate the normality of data distribution and determine whether

parametric or nonparametric tests should be used to evaluate the data. According to the obtained results, the distribution of data is normal, because for all of the variables sig>0.05. Then, it is possible to use parametric tests to analyse data.

The first research question sought if there is a significant relationship between EFL learners' scaffolding via Google Docs and writing achievement. The result is shown in Table 1 using Pearson's correlation coefficient.

Table 1. Relationship between EFL learners' scaffolding via Google Docs and writing achievement

		Writing achievement	Instructional	Interactional	Reflective	Affective	Total
Writing achievement	Pearson Correlation	1					
	Sig.						
Instructional	Pearson Correlation	.065	1				
	Sig. (2-tailed)	.544					
Interactional	Pearson Correlation	.159	.568**	1			
	Sig. (2-tailed)	.135	.000				
Reflective	Pearson Correlation	.093	.473**	.533**	1		
	Sig. (2-tailed)	.385	.000	.000			
Affective	Pearson Correlation	.305**	.380**	.359**	.397**	1	
	Sig. (2-tailed)	.003	.000	.001	.000		
Total	Pearson Correlation	.203	.812**	.729**	.803**	.600**	1
	Sig. (2-tailed)	.055	.000	.000	.000	.000	

**. Correlation is significant at the 0.01 level (2-tailed).

As it can be seen in Table 1, all subscales of EFL learners' scaffolding via Google Docs and its overall score were not significantly related to writing achievement. Therefore, the hypothesis of the research was rejected, and the opposite hypothesis was confirmed.

The second research question sought whether there is a significant relationship between EFL learners' scaffolding via Google Docs and EFL learners' writing metacognitive knowledge. The result can be found in Table 2 using Pearson's correlation coefficient.

Table 2. EFL learners' scaffolding via Google Docs and EFL learners' writing metacognitive knowledge

		Writing metacognitive awareness	Instructional	Interactional	Reflective	Affective	Total
Writing metacognitive awareness	Pearson Correlation	1					
	Sig. (2-tailed)						
Instructional	Pearson Correlation	.652**	1				
	Sig. (2-tailed)	.000					
Interactional	Pearson Correlation	.714**	.568**	1			
	Sig. (2-tailed)	.000	.000				
Reflective	Pearson Correlation	.593**	.473**	.533**	1		
	Sig. (2-tailed)	.000	.000	.000			
Affective	Pearson Correlation	.550**	.380**	.359**	.397**	1	
	Sig. (2-tailed)	.000	.000	.001	.000		
Total	Pearson Correlation	.826**	.812**	.729**	.803**	.600**	1
	Sig. (2-tailed)	.000	.000	.000	.000	.000	

**. Correlation is significant at the 0.01 level (2-tailed).

As can be seen in Table 2, all subscales of EFL learners' scaffolding via Google Docs including instructional, interactional, reflective, affective, and also its overall score were significantly related to writing metacognitive knowledge. Therefore, the research hypothesis was confirmed, and the counter hypothesis was rejected. Multiple regression was used for further analysis. The results of this test are shown in Tables 4-5.

Table 3. Model Summary of EFL learners' scaffolding and their writing metacognitive knowledge

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.825[a]	.681	.666	.31923

As seen in Table 3, there are a variety of indicators that evaluate the success of the model in predicting the dependent variable. Here, the R value obtained is equal to 0.825. That is the Pearson correlation between the values of instructional, interactional, reflective, and affective and the actual values of the writing metacognitive knowledge. The R^2 coefficient shows the amount of variance explained by the writing metacognitive knowledge by the combination of the instructional, interactional, reflective, and affective which is 0.681. In other words, these four variables together explain 68% of the variance of the writing metacognitive knowledge variable. Other indicators such as Adjusted R Square (0.666) and Std. Error of the Estimate (0.31923) can be found in the Table.

Table 4. ANOVA for EFL learners' scaffolding and their writing metacognitive knowledge

	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	18.475	4	4.619	45.324	.000
	Residual	8.662	85	.102		
	Total	27.137	89			

Based on Table 4, the F value observed in the 4 degrees of freedom is equal to 45.324 and P=<0.05 shows that this F value is significant at the 0.05 level. Therefore, the variance of variable writing metacognitive knowledge can be explained by variables instructional, interactional, reflective, and affective. Table 5 shows the prediction coefficients of writing metacognitive knowledge using these predictor variables.

Table 5. Coefficients (predictor variables of EFL learners' writing metacognitive knowledge model)

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	.519	.178		2.906	.005
	Instructional	.179	.054	.258	3.302	.001
	Interactional	.311	.064	.392	4.874	.000
	Reflective	.130	.060	.165	2.156	.034
	Affective	.156	.044	.246	3.561	.001

As illustrated in Table 5, four independent predictor variables namely instructional, interactional, reflective, and affective are related to writing metacognitive knowledge P < 0.005. Therefore, based on the alpha level of 0.05, they can statistically explain the variance of writing metacognitive knowledge in a meaningful way. The standardized beta (β) coefficient shows the influence coefficient of the instructional (β=0.258) and according to the t-statistic (3.302) it can be inferred that this variable can reflect the changes regarding writing metacognitive knowledge. This coefficient of influence is positive, and it shows that if one unit is added to the amount of instructional, the score of writing metacognitive knowledge increases by 25.8%. The standardized beta (β) coefficients show the influence coefficient of the interactional (β=0.392) and according to the t-statistic (4.874) it can be inferred that this variable can reflect the changes related to writing metacognitive knowledge. This coefficient is positive, and it shows that if one unit is added to the amount of instructional, the score of writing metacognitive knowledge increases by 39.28%. The standardized beta (β) coefficients show the influence coefficient of the reflective (β=0.165) and according to the t-statistic (2.156) it can be inferred that this variable can reflect the changes related to writing metacognitive knowledge. This coefficient of influence is positive, and it shows that if one unit is added to the amount of Instructional, the score of writing metacognitive knowledge increases by 16.5 %. The standardized beta (β) coefficients show the influence coefficient of the affective (β=0.246) and according to the t-statistic (3.561) it can be inferred that this variable can reflect the changes related to writing metacognitive knowledge. This coefficient of influence is positive, and it shows that if one unit is added to the amount of instructional, the score of writing metacognitive knowledge increases by 24.6%.

Research question three explored if there is a significant relationship between EFL learners' writing metacognitive knowledge and their writing achievement. The result of Pearson's correlation coefficient is presented in Table 6.

Table 6. Relationship between EFL learners' writing metacognitive knowledge and writing achievement

		writing achievement	writing metacognitive knowledge
writing achievement	Pearson Correlation	1	.275**
	Sig. (2-tailed)		.009
writing metacognitive knowledge	Pearson Correlation	.275**	1
	Sig. (2-tailed)	.009	

**. Correlation is significant at the 0.01 level (2-tailed).

As can be seen in Table 6, writing achievement was significantly related to writing achievement. Therefore, the research hypothesis was confirmed, and the counter hypothesis was rejected. Multiple regression was used for further analysis. The results of this test are illustrated in Tables 7-9.

Table 7. Model Summary of EFL learners' writing metacognitive knowledge and writing achievement

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.275[a]	.076	.065	.59319

As displayed in Table 7, there are a variety of indicators that evaluate the success of the model in predicting the dependent variable. Here, the R value obtained is equal to 0.275. That is the Pearson correlation between the values of writing metacognitive knowledge and the actual values of the writing achievement. The R^2 coefficient shows the amount of variance explained by the writing achievement by the writing metacognitive knowledge which is 0.076 here. In other words, these four variables together explain 7.6% of the variance of the writing achievement variable. Other indicators such as Adjusted R Square (0.065) and Std. Error of the Estimate (0.59319) can be seen in the Table.

Table 8. ANOVA for EFL learners' writing metacognitive knowledge and their writing achievement

	Model	Sum of Squares	df	Mean Square	F	Sig.
1	Regression	2.543	1	2.543	7.227	.009
	Residual	30.965	88	.352		
	Total	33.508	89			

According to Table 8, the F value observed in the 1 degree of freedom is equal to 7.227 and P=0.000<0.05 indicates that this F value is significant at the 0.05 level. Therefore, the variance of variable writing achievement can be explained by the writing metacognitive knowledge variable. Table 10 shows the prediction coefficients of writing metacognitive knowledge using these predictor variables.

Table 9. Coefficients (predictors of variables of EFL learners' writing achievement

		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
1	(Constant)	2.226	.331		6.719	.000
	writing metacognitive knowledge	.306	.114	.275	2.688	.009

As illustrated in Table 9, one independent predictor variable namely writing metacognitive knowledge is related to writing achievement P<0.005. Therefore, based on the alpha level of 0.05, they can statistically explain the variance of writing achievement in a meaningful way. The standardized beta (β) coefficient shows the influence coefficient of the writing metacognitive knowledge (β=0.275) and according to the t-statistic (2.688) it can be inferred that this variable can reflect the changes in writing achievement. This coefficient of influence is positive, and it shows that if one unit is added to the amount of writing metacognitive knowledge, writing achievement increases by 27.5%.

Since it is not possible to investigate the role of the mediator dependent variable using regression, it was decided to investigate the primary model and the mediator role of writing metacognitive knowledge using path analysis.

Research question four investigated if EFL learners' writing metacognitive knowledge mediates the relationship between scaffolding via Google Docs and their writing achievement. To answer the question, path analysis was employed.

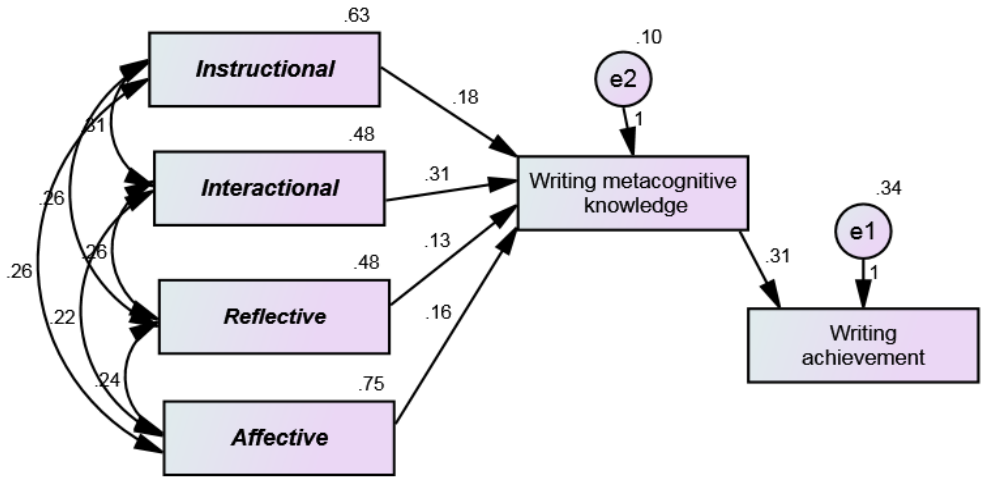

Figure 3. The final model

For this purpose, three categories of Absolute Fit Indices, Comparative Fit Indices and Parsimonious Fit Indices were employed in examining the relationship model between these variables, using Amos software. The results related to some of the most important indicators of absolute, comparative, and parsimonious fit are presented in Table 10.

Table 10. Model fit indices

Index name	Acceptable fit	Model evaluation result
CMIN (χ^{r})	-	6.616
CMIN/DF	<3	1.654 (P=.158)
Goodness of Fit Index (GFI)	>.90	. 977
Adjusted Goodness of Fit Index (AGFI)	>.90	.878
Normed Fit Index (NFI)	>.90	.968
Comparative Fit Index (CFI)	>.90	.986
Incremental Fit Index (IFI)	>.90	.987
Root Mean square Residual (RMR)	<.05	.023
Root Mean Square Error of approximation (RMSEA)	<.05	.086

Based on Table 10, the relative chi-square is equal to 1.654, which means that the model does not have a good fit. The goodness of fit index (GFI) jointly evaluates the relative value of variances and covariance through the model. Here, the GFI value is equal to 0.977 and indicates the good fit of the model. In addition, in the present study, the Root Mean Square Residual (RMR) is equal to 0.023, which indicates a good fit, and as a result, the model is very desirable. Normalized fit index (NFI) is acceptable for values above 0.90. This index in the current model is equal to 0.968, which shows the optimal fit of the model. The value of Comparative Fit Index (CFI) is equal to 0.986, which indicates the fit of the model. The Incremental Fit Index (IFI) was developed to address parsimony and sample size issues related to NFI. This index here is equal to 0.987. The Root Mean Square Error of approximation (RMSEA) represents the average of the differences between the actual correlation/variance of the sample and the expected model estimated from the population. Considering that the current model, RMSEA=0.086>0.08, does not have a good fit, In general, the result of the indicators reveals that all the indicators have been reported in the RMSEA component at the optimal level, and the model has a good fit with the data, and this indicates that there is a linear relationship between the variables and the structure.

Although the presented model has almost a good fit, the question arises as to what is the significant effect of the variables instructional, interactional, reflective, and affective on the variable writing achievement through variable writing metacognitive knowledge as the dependent variable of the criterion. The results are presented in Table 11.

Table 11. The direct and indirect effect of independent variables with the dependent medium and criterion

Predictor	Criterion	Direct effects	Indirect Effect	Total Effect
Instructional	writing metacognitive knowledge	0.179	0.000	0.179
Interactional	writing metacognitive knowledge	0.311	0.000	0.311
Reflective	writing metacognitive knowledge	0.130	0.000	0.130
Affective	writing metacognitive knowledge	0.156	0.000	0.156
Writing metacognitive knowledge	writing metacognitive	0.306	0.000	0.306

As it can be seen in Table 12, the results indicate that writing metacognitive knowledge has a mediating effect on the relationship between instructional, interactional, reflective, and affective with writing metacognitive.

5. Discussion

The aim of the current investigation was to test the hypothesis that writing metacognitive knowledge mediates the relationship between scaffolding via Google Docs and academic writing achievement. As a result, the first research question explored if there was a significant relationship between EFL learners' scaffolding via Google Docs and EFL learners' writing achievement. The results demonstrated that EFL learners' scaffolding via Google Docs as well as any of its subscales did not significantly correlate with writing achievement. Based on the results, the researchers concluded that there was no significant relationship between the scaffolding of EFL learners via Google Docs and their writing achievement. The results of the first research question were compatible with those of Nhung and Hue (2022), who concluded that use of Google Docs had a significant effect on EFL students' writing performance. In the same line, Nguyen and Nguyen (2022) reported that applying Google Docs in online teaching and learning English had positive results in students' writing performance. Seyyedrezaie et al. (2016) also investigated the impact of the Google Docs on EFL learners' writing achievement and found that Google Docs improved learners' writing performance statistically. In the same vein, Fathi et al. (2021) who investigated the impact of collaborative writing using Google Docs on EFL learners' writing performance and writing self-regulation found that collaborative writing both via using Google Docs and in the face-to-face classroom significantly promoted the writing performance and writing self-regulation of the participants. This finding of the first research question is not unexpected because scaffolding via Google Docs may provide a dynamic and interactive platform that supports real-time feedback, peer collaboration, and structured guidance which play a key role in improving writing achievement. The collaborative nature of Google Docs allows learners to engage in meaningful interactions, share ideas, and refine their writing through continuous revision. This supports Vygotsky's (1978) sociocultural theory of learning. According to this theory, learning is a socially mediated process, and tools like Google Docs facilitate scaffolding by enabling more knowledgeable peers to support learners within their zone of proximal development (ZPD). In addition, the accessibility and flexibility of Google Docs may encourage active participation and self-regulation.

The second research question explored if there was a significant relationship between scaffolding via Google Docs and EFL learners' writing metacognitive knowledge. The results uncovered a strong positive relationship between scaffolding via Google Docs and EFL learners' writing metacognitive knowledge, indicating that EFL learners' scaffolding via Google Docs is connected with their level of writing metacognitive knowledge. This can display how EFL learners' scaffolding via Google Docs is crucial to foster their learning autonomy and self-regulation and to be an autonomous learner who is able to regulate his/her own thinking, to be more aware of the significance of writing metacognitive knowledge which develops his/her thinking skills, and to promote students' achievement in online learning. The result lend support to the study conducted by Mortazavi et al. (2016) who reported that scaffolding strategies develop the student's self-efficacy and monitoring skills during writing. They also declared that notable developments in self-regulatory and writing skills are the results of applying scaffolding mechanisms. In a partially similar result, Ikwati (2020) also investigated scaffolding in teaching writing and revealed that through scaffolding not only do students learn how to do the assignment but also it enables them to become independent and successful writers. That finding is not unexpected since the aim of scaffolding via Google Docs is to transfer learning responsibility, encourage learners' autonomy, foster students' success in learning, gain self-regulation, and regulate their own thinking. Moreover, metacognition, an essential segment of self-regulation (Teng, 2019), is of great significance in online learning contexts (Ersani et al., 2021).

The third research question aimed to find if there was a significant relationship between writing metacognitive knowledge and academic writing achievement. This is compatible with the finding that found that there is a positive relationship between writing metacognitive knowledge and ESL learners' writing performance (Kasper, 1997). This is also in tandem with the outcomes revealed by Victori (1999) who studied the relationship between L2 learners' metacognitive abilities and writing performance and perceived a meaningful relationship between the two variables. The findings of the present study also support the study by Teng (2019) who studied the relationship between metacognitive awareness and writing achievement. He found that metacognitive awareness is an essential segment in exhibiting an EFL learner's writing achievement and that metacognitive regulation can process a learner's metacognitive awareness and promote independence in developing desired written outcomes. Furthermore, the result is consistent with Colognesi et al. (2020) who reported those FL/SL students with higher metacognitive knowledge made more significant progress in their writing skill.

The fourth research question explored the mediating role of writing metacognitive knowledge. The results revealed that the relationship between EFL learners' scaffolding via Google Docs and their academic writing achievement through writing metacognitive knowledge was certainly noteworthy. In other words, writing metacognitive knowledge gained a mediating role in the relationship between scaffolding via Google Docs and academic writing achievement. This means that EFL learners who employ scaffolding via Google Docs tend to have significant writing metacognitive knowledge and ultimately may have special academic writing achievement. This is considered reasonable because if learners do not employ scaffolding via Google Docs, sources of enhancing purposeful communication, reciprocal teaching, autonomous class participation, self-regulation, dynamic classroom interaction, mutual collaboration, as well as the transfer of learning responsibility could diminish and in turn learners' academic writing may be influenced. Despite the fact that to the researchers' knowledge, no studies have specifically explored the precise relationship between learners' scaffolding via Google Docs and their academic writing achievement through writing metacognitive knowledge, some studies have demonstrated a relationship between learners' scaffolding via Google Docs and their writing metacognitive knowledge (Ersani et al., 2021; Ikawati, 2020; Mortazavi et al., 2016) and writing metacognitive

knowledge and their academic writing achievement (Al-Jarrah et al., 2018; Colognesi et al., 2020; Ramadhanti & Yanda, 2021; Teng, 2019; Wijaya, 2022).

One acceptable interpretation for the mediating effect of writing metacognitive knowledge could be that "metacognitive awareness includes what learners know about their own thinking and what they know about strategies for learning" (Stanton, et al., 2021, P. 2). Writing metacognitive awareness helps learners identify their own strengths (Ramadhanti & Yanda, 2021) and strategies (Jaleel & Premachandran, 2016), requires learners to reflect on what they write (Colognesi, et al., 2020), and helps learners develop an awareness of writing process (Hayes & Flower, 1980). Thus create adept writers who are able to regulate their own thinking (Jaleel & Premachandran, 2016) and learning (Yanyan, 2010).

6. Conclusion

The results illuminated the considerable effect of the mediation of writing metacognitive knowledge in the relationship of scaffolding via Google Docs and academic writing achievement among EFL learners. Thereupon, raising the level of one of these variables has an impact on the development of other variables. It follows that to improve EFL learners' academic writing achievement, it is essential, first, to reinforce their writing metacognitive knowledge which is subsequently impacted by their scaffolding via Google Docs. The results of this study have substantial implications for EFL instructors and curriculum developers, because it reinforces the role of metacognitive knowledge in the domain of foreign language writing. The findings of this study have also implicated that raising metacognitive knowledge requires serious attention in teacher training courses. Finally, the present study adds to the body of studies in the field of metacognition and contributes to the understanding of writing metacognitive knowledge.

Scaffolding via Google Docs can be effectively integrated into metacognitive teaching methods in EFL writing programs, where appropriate training and activities can be offered in the classroom setting. Educators should focus on enhancing students' metacognitive understanding and strategies while also leveraging technology to broaden EFL learners' range of metacognitive experiences, thereby enhancing their writing skills. In this regard, EFL teachers could improve EFL students' metacognitive experiences and writing skill by encouraging them to use Google Docs inside and outside classrooms.

It is worth noting that the present study has some limitations. recruiting EFL learners from a province in the West of Iran is the study's first drawback. As a result, they are not representative of all English language learners from all provinces of the country. Future investigations may utilize the questionnaires among EFL learners selected from a wide range of institutions throughout the country. By the same token, as the participants lacked any prior experience of using Google Docs for writing development, learners' writing performance and their self-regulation might have been affected by a kind of Hawthorne effect experienced during the course. Additionally, the variables were only evaluated via self-reporting. Future research could employ techniques such as integrating concurrent think-aloud protocols, eye tracking, recordings of learner-system interactions on screen, log files, and other similar methods. (Azevedo, 2020). In further research, the researcher might require thinking about performing in-depth studies with a variety of data collection techniques (e.g. think loud, interviews, journals) and drawing on additional facets (e.g. teachers' attitudes) in order to provide more reliable results.

7. References

Al-Jarrah, T.M., Mansor, N., Talafhah, R.H., & AL-Jarrah, J.M. (2018). Instructions based on metacognitive writing module on 12th grade students' writing achievements. *International Journal of Current, 10*(12), 76773-76779. doi: https://doi.org/10.24941/ ijcr.33626.12.2018.

Apple, K.J., Reis-Bergan, M., Adams, A.H., & Saunders, G. (2011). Online tools to promote student collaboration. In D.S. Dunn, J.H. Wilson, J. Freeman, & J.R. Stowell (Eds.), *Getting connected: Best practices for technology enhanced teaching and learning in high education* (pp. 239-252). New York, NY: Oxford University Press.

Avarzamani, F. & Farahian, M. (2019). An investigation into EFL learners' reflection in writing and the inhibitors to their reflection. *Cogent Psychology*, 1-13. https://doi.org/10.1080/23311908.2019.1690817

Azevedo, R. (2020). Reflections on the field of metacognition: issues, challenges, and opportunities. *Metacognition and Learning,* 91-98. https://doi.org/10.1007/s11409-020-09231-x

Blair, R.M. (2006). *Beyond crossroads: Implementing mathematics standards in the first two years of college*. Memphis, TN: American Mathematical Association of Two-Year Colleges.

Bottge, B., Rueda, E., Kwon, J., Grant, T., & LaRoque, P. (2009). Assessing and tracking students' problem-solving performances in anchored learning environments. *Educational Technology Research and Development, 57(4),* 529-552. https://doi.org/10.1007/s11423-007-9069-y

Cai, Y., Zhao, C. (2023). Metacognitive strategies and self-efficacy co-shape L2 achievement: A multilevel structural equation modeling approach. *System, 117*. https://doi.org/10.1016/j.system.2023.103099

Cheraghi, Z., Nejati, R., & Bakhtiari (2022). The relationship between Iranian EFL learners' critical thinking disposition and their writing strategy use. *Iranian Journal of Learning and Memory, 5(*17), 51-58. https://doi.org/10.22034/iepa.2022.157661

Cho, K., & MacArthur, C. (2010). Student revision with peer and expert reviewing. *Learning and Instruction, 20*(4), 328–338. https://doi.org/10.1016/j.learninstruc.2009.08.006

Colognesi, S., Piret, C., Demorsy, S., & Babier, E. (2020). Teaching writing- with or without metacognition? an exploratory study of 11-to 12-year-old students writing a book Review. *IEJEE International Electronic Journal of Elementary Education, 12*(5), 459-470. https://doi.org/10.26822/iejee.2020562136

Darwin, W., Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2023). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education, 11*(1). https://doi.org/10.1080/2331186X.2023.2290342

Dennis, J.L., Somerville, M.P. (2022). Supporting thinking about thinking: examining the metacognition theory-practice gap in higher education. High Educ. https://doi.org/10.1007/s10734-022-00904-x

Dizon, G. (2016). A comparative study of Facebook vs. paper -and -pencil writing to improve L2 writing skills. *Computer Assisted Language Learning, 29*(8), 1249 – 1258. https://doi.org/10.1080/09588221.2016.1266369

Ebadi, S., & Rahimi, M. (2017). Exploring the impact of online peer-editing using Google Docs on EFL learners' academic writing skills: A mixed methods study. *Computer Assisted Language Learning, 30*(8), 787–815. https://doi.org/10.1080/09588221.2017.1363056

Eriyani, E. (2020). Metacognition awareness and its correlation with academic achievement of student teachers: The Case of One Private Higher Education Institution. *IRJE Indonesian Research Journal in Education, 4*(1), 78-88. https://doi.org/10.22437/irje.v4i1.8573

Ersani, N.P.D., Suwastini, N.K.A., & Artini, L.P. (2021). Schemes of scaffolding in online education. *Retorika: Journal Ilmu Bahasa, 7*(1), 10-18. doi: https://doi.org/10.22225/ jr.7.1.2941.10-18.

Farahian, M. (2017). Developing and validating a metacognitive writing questionnaire for EFL learners. *Issues in Educational Research, 27*(4), 736-750.

Farahian, M., & Avarzamani, F. (2018). The impact of portfolio on EFL learners' metacognition and writing performance. *Cogent Education*, 1-21. https://doi.org/10.1080/2331186x.2018.145091

Farahian, M., & Ebadi, S. (2022). Collaborative digital writing and metacognitive knowledge in writing among TEFL students: the mediating role of online knowledge sharing. *Research and Practice in Technology Enhanced Learning*, 18. Retrieved from https://rptel.apsce.net/index.php/RPTEL/article/view/2023-18005

Fathi, J., Arabani,A.S.,& Mohamad, P. (2021). The Effect of collaborative writing using Google Docs on EFL learners' writing performance and writing self-regulation. *Language Related Research, 12*(5), 333- 359. http://dx.doi.org/10.52547/LRR.12.5.13

Flavell, J. H. (1987). Speculations about the nature and development of metacognition. In Metacognition, motivation, and understanding, edited by F. Weinert and R. Kluwe, 9-21. Erlbaum.

Flavell, J.H. (1979). Metacognition and cognitive monitoring: A new area of cognitive-.developmental inquiry. *American Psychologist, 34*(10), 906911. http://dx.doi.org/10. 1037/0003-066X.34.10.906

Flower, L., Wallace, D. L., Norris, L., & Burnett, R. A. (1994). *Making thinking visible: Writing, collaborative planning, and classroom inquiry*. National Council of Teachers of English.

Geng, F.,Yu, S., Liu, C., & Liu, Z. (2021). Teaching and learning writing in English as a foreign language (EFL) School education contexts: A thematic review. *Scandinavian Journal of Educational Research,66*(3), 491-504. https://doi.org/10.1080/00313831.2021.1897872

Graham, S., & Perin, D. (2007). *Writing next: Effective strategies to improve writing of adolescents in middle and high schools*. Alliance for Excellent Education.

Hayes, J. & Flower, L. (1980). Identifying the organization of writing processes. In L. Gregg & E. Steinberg (Eds.), *Cognitive processes in writing* 3-30. Lawrence Erlbaum.

Hemati, Z., & Farahian, M. (2024). Fostering EFL Learners' Classroom Engagement: Does Collaborative Writing through Google Docs Make a Difference?. *Technology Assisted Language Education*, *2*(1), 1-20. https://doi.org/10.22126/tale.2024.10225.1030

Hogan, K., & Pressley, M. (1997). *Scaffolding student learning: Instructional approaches and issues*. Brookline Books.

Ikawati, L. (2020). Scaffolding in Teaching Writing. *Al-Tarbiyah, 30*(1), 48-58. https://doi.org/10.24235/ath.v%vi%i.6487

Ilgun Dibek, M., Sahin Kursad, M., & Erdogan, T. (2024). Influence of artificial intelligence tools on higher order thinking skills: a meta-analysis. *Interactive Learning Environments*, 1–23. https://doi.org/10.1080/10494820.2024.2402028

Jacobs, H.L., Zinkgraf, S.A., Wormuth, D.R., Hartifiel, V.G., & Haughey, J.B. (1981). Testing ESL composition: A practical approach. Rowley, MA: Newbury House Publishers. http://dx.doi.org/10.1080/00461520.1987.9653052.

Jaleel, S & Premachandran. P. (2016). A study on the metacognitive awareness of Secondary School Students. *Universal Journal of Educational Research, 4*(1), 165-172. https://doi.org/10.13189/ujer.2016.040121

Jones, R.W. (2007). Learning and teaching in small groups: Characteristics, benefits, problems and approaches. *Anaesthesia and Intensive Care, 35*(4), 587-592. https://doi.org/10.1177/0310057X0703500420

Kadkhodaei, S., Talebinejad, M. R. and Shahrokhi, M. (2025). Predicting Teachers' Stroke, through Identity, Immunity, and TPACK: Iranian EFL Teachers in Focus. *Applied Linguistics Inquiry*, *3*(1), 133-146. doi: 10.22077/ali.2025.8873.1090

Kasper, L.F. (1997). Assessing the metacognitive growth of ESL student writers. *TESL-E 3*(1), 1-20. http://www.tesl-ej.org/wordpress/issues/volume3/ej09/ej09a1.

Khotimah, K., Rusijono, Mariono, A. (2024). Enhancing metacognitive and creativity skills through AI-driven meta-learning strategies. *International Journal of Interactive Mobile Technologies, 18*(5), pp. 18–31. https://doi.org/10.3991/ijim.v18i05.47705

Maftoon, P., Birjandi, P. & Farahian, M. (2014). Investigating Iranian EFL learners' writing metacognitive awareness. *International Journal of Research Studies in Education, 3*(5), 37-52. https://doi.org/10.5861/ijrse.2014.896.

Mahmood, N. (2018). *Using Google Docs to enhance collaborative writing skills among non-native English-speaking students*. Retrieved from https://www.learntechlib.org/primary/p/178319/

McCormick, C.B. (2003). Metacognition and learning. In Handbook of Psychology, Vol. 7, *Educational Psychology*, edited by William M. Reynolds and Gloria E. Miller, 79–102. New Jersey: John Wiley& Sons.

McCutchen, D. (2000). Knowledge acquisition, processing efficiency, and working memory: Implications for a theory of writing. *Educational Psychologist, 35*, 13–23. https://doi.org/10.1207/S15326985EP3501_3

Mohammadzadeh, A., Ahour, T., & Saeidi, M. (2022). Investigating Iranian EFL learners' negotiation types in writing with different scaffolding patterns. Journal of Language *Horizons, 6*(1), 95-110. https://doi.org/10.22051/lghor.2020.30357.1259

Mortazavi, M., Jafarigohar, M., Rouhi, A., & Soleimani, H. (2016). The effect of scaffolding through structuring and problematizing on EFL Learners' writing self-regulatory skills, essay writing skill and global planning time. *Quarterly Journal of Research in School and Virtual Learning, 3*(15),15-56.

Nabhan, S., & Sa'diyah, H. (2021). Collaborative writing using Google Docs in an EFL classroom: Voices from high school students. *VOLES Voices of English Language Education Society, 5*(2), 156-166 https://dx.doi.org/10.29408/veles.v5i2.3863.

Nhung, N.T.H., & Hue, N.T.T. (2022). Use of Google Docs in teaching and learning English online to improve students' writing performance. *International Journal of TESOL & Education, 2*(2), 186–200. https://doi.org/10.54855/ijte.222210.

Nosratinia, M., & Adibifar, Sh. (2014). The effect of teaching metacognitive strategies on Field-dependent and Independent learners' writing. *Procedia - Social and Behavioral Sciences, 98*, 1390–1399. https://doi.org/10.1016/j.sbspro.2014.03.557

Nourazar, S., Kakvand, R., & Aliasin, S.H. (2022). The impact of scaffolded metacognitive writing strategy instruction on Iranian intermediate EFL learners' IELTS writing task 2. Education Research International. https://doi.org/10.1155/2022/6297895

Owens, R.E. (2012). *Language Development: An introduction* (8th ed.). Pearson Education, Inc., Upper Saddle River. https://doi.org/013258252X.

Perron, B., & Sellers, J. (2011). A review of the collaborative and sharing aspects of Google Docs. *Research on Social Work Practice, 21*, 489-490. https://doi.org/10.1177/1049731510391676

Ramadhanti, D., & Yanda, D. P. (2021). Students' metacognitive awareness and its impact on writing skill. *International Journal of Language Education, 5*(3), 193-206. https://doi.org/10.26858/ijole. v5i3.18987

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review, 7*(4), 351-371. https://doi.org/10.1007/BF02212307

Semeraro, J., & Moore, N. S. (2016). The use of Google docs technology to support peer revision. *Writing Instruction to Support Literacy Success, 203* –220. https://doi.org/10.1108/S2048-045820160000007013

Seyyedrezaie, Z. S., Ghonsooly, B., Shahriari, H., & Fatemi, H. H. (2016). Mixed methods analysis of the effect of Google Docs environment on EFL Learners' writing. performance and causal attributions for success and failure. *Turkish Online Journal of Distance Education (TOJDE), 17*, 90–110. https://doi.org/10.17718/tojde.34418

Slavin, R.E. (1996). Research on cooperative learning and achievement: What we know, what we need to write. *Contemporary Educational Psychology, 21*, 43–69. http://dx.doi.org/10.1006/ceps.1996.0004

Stanton, J. D., Sebesta, A. J., & Dunlosky, J. (2021). Fostering metacognition to support student learning and performance. *CBE-Life Sciences Education, 20*(3), 1-7. https://doi.org/10.1187/cbe.20-12-0289.

Stone, C.A. (1998). The metaphor of scaffolding: its utility for the field of learning disabilities. *Journal of Learning Disabilities, 31*, 344–364. https://doi.org/10.1177/002221949803100404

Strobl, C. (2013). Affordances of Web 2.0 technologies for collaborative advanced writing in a foreign language. *Calico Journal, 31*(1), 1–18.

Sun, Q., & Zhang, L.J. (2022). Understanding learners' metacognitive experiences in learning to write in English as a foreign language: A structural equation modeling approach. *Front. Psychol*. https://doi.org/10.3389/fpsyg.2022.986301

Suwastini, N. K. A., Ersani, N. P. D., Padmadewi, N. N., & Artini, L. P. (2021). Schemes of Scaffolding in Online Education. *Retorika: Jurnal Ilmu Bahasa, 7*(1), 10–18. https://doi.org/10.22225/jr.7.1.2941.10-18

Szmyd, K., & Mitera, E. (2024). The Impact of artificial intelligence on the development of critical thinking skills in students. *European Research Studies Journal, XXVII* (2), 1022-1039. https://doi.org/10.35808/ersj/3876

Teng, M. F. (2021). Interactive-whiteboard-technology-supported collaborative writing: Writing achievement, metacognitive activities, and co-regulation patterns. *System, 97*, 102426. https://doi.org/10.1016/j.system.2020.102426

Teng, M.F. (2019). The role of metacognitive knowledge and regulation in mediating university EFL learners' writing performance. *Innovation in Language Learning and Teaching*, 1–15. https://doi.org/10.1080/17501229.2019.1615493

Vakilifard, A., Abedini, A. (2021). Investigating the employment of metacognitive strategies in listening comprehension among non-Iranian language learners. *Journal of English Language Teaching and Learning, 13*(28), 261-277. https://doi.org/10.22034/elt.2021.48515.2455

Van de Pol, J., Volman, M., & Beishuizen, J. (2010). Scaffolding in teacher-student interaction: A decade of research. *Educational Psychology Review*, 22, 271-296. https://doi.org/10.1007/s10648-010-9127-6

Victori, M. (1999). An analysis of writing knowledge in EFL composing: A case study of two effective and two less effective writers. *System, 27*(4), 53 555. https://doi.org/10.1016/S0346-251X(99)00049-4

Vygotsky, L.S. (1978). *Mind in society.* Cambridge: Harvard University Press.

Vygotsky, L.S. (1986). Thought and language. MIT Press.

Wang, K., Zhang, L.J. & Cooper, M. (2024). Metacognitive instruction for improving the effectiveness of collaborative writing for EFL learners' writing development. *Asia-Pacific Edu Res*. https://doi.org/10.1007/s40299-024-00886-7

Warschauer, M. (2011). Learning in the cloud: *How (and why) to transform schools with digital media*. Teachers College Press.

Widyastanti, M. (2019). Using Google Docs on collaborative writing activity in SMPN1 Driyorejo-Gresik. *Digital. Library*. 1-58.

Wijaya, K.F. (2022). Exploring English education Master Students' metacognitive awareness in academic writing learning enterprises. *ELTR Journal, 6*(1), 11-25. https://doi.org/10.37147/eltr.2022.060102

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem-solving. *Journal of Child Psychology and Psychiatry and Allied Disciplines, 17*, 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Yanyan, Z. (2010). Investigating the role of metacognitive knowledge in English writing. *HKBU Papers in Applied Language Studies, 14*, 2546.

Yu-Ling, Y., & Shih-Guey, J. (2001). *Investigating the metacognitive awareness and strategies of English-majored university student writers.*

Zarinkamar, N., Abdi, R., & Davaribina, M. (2021). Developing EFL learners' creativity in writing skill through dynamic assessment-based teaching. *Research in English Language Pedagogy, 9*(2), 309-338. https://doi.org/10.30486/relp.2021.1914150.1237

Zheng, B., Lawrence, J., Warschauer, M., & Lin, C. H. (2015). Middle school students' writing and feedback in a cloud-based classroom environment. Technology, Knowledge and Learning, 20(2), 201–229. https://doi.org/10.1007/s10758-014-9239-z

Zhou, W., Simpson, E., & Domizi, D. P. (2012). Google Docs in an out-of-class collaborative writing activity. International Journal of Teaching and Learning in Higher Education, 24(3), 359-375.

Original Article

University of Birjand

Applied Linguistics Inquiry

Spring 2025, Vol. 3, No. 1, pp. 162-164

Applied Linguistics Inquiry

https://doi.org/10.22077/ali.2025.8602.1059

Review of AI in Language Teaching, Learning, and Assessment (2024) edited by Fang Pan. IGI Global: 384 pp., ISBN: 9798369308738

Neda Kianinezhad[1*]

[*1]Ph.D. Candidate in Applied Linguistics, Department of English Language, Faculty of Literature and Humanities, Hakim Sabzevari

ARTICLE HISTORY

Received: 18 December 2024
Revised: 06 January 2025
Accepted: 03 February 2025
Published: 31 March 2025

CORRESPONDING AUTHOR

E-mail: neda.kianinezhad@hsu.ac.ir

ABSTRACT

AI in Language Teaching, Learning, and Assessment (2024), edited by Fang Pan, offers a comprehensive and insightful exploration of how artificial intelligence (AI) is reshaping language education. The book is thoughtfully divided into three well-structured sections Preview, Perspectives, and Practice, each addressing essential dimensions of AI integration. In the Preview section, Pan examines AI's transformative role in enhancing educational delivery and personalizing learning experiences, paving the way for greater efficiency and accessibility. Subsequently, the Perspectives section collates diverse global viewpoints that critically address challenges such as algorithmic bias, data privacy, and equity concerns. Finally, the Practice section offers evidence-based case studies, showcasing tangible innovations such as adaptive tutoring systems and AI-powered assessment tools, which have demonstrated real-world impact. More importantly, Pan strikes an impressive balance between highlighting AI's vast potential and scrutinizing its ethical and practical complexities. By combining theoretical perspectives, empirical findings, and practical applications, this book stands as an essential resource for educators, researchers, and policymakers, aiming to navigate AI-driven developments in language education with both responsibility and clarity.

KEYWORDS: Language teaching; Language learning; Assessment tools; Personalized learning

1. Introduction

In today's educational landscape, integrating technology into language teaching is essential. It enhances traditional methods and introduces innovative strategies to meet diverse learning needs. By leveraging digital tools, language education becomes more interactive, personalized, and accessible (Kianinezhad, 2023). To succeed in the digital age, both educators and students must effectively use technology to create a dynamic and impactful learning experience. Accordingly, in *AI in Language Teaching, Learning, and Assessment*, Fang Pan delivers a timely and rigorous analysis of the transformative role artificial intelligence (AI) is playing in language education. At a time when technological innovations are rapidly gaining momentum, this book bridges the crucial gap between research, policy formation, and practical implementation. Besides, Pan effectively demonstrates how AI-driven solutions are enhancing pedagogical practices, fostering learner engagement, increasing assessment precision, and broadening educational accessibility on a global scale.

Importantly, the volume does not merely emphasize the positive potential of AI; it also highlights critical socio-educational and ethical considerations that arise alongside its adoption. This balanced approach ensures that the book serves as an essential guide for educators, policymakers, and researchers alike, empowering them to leverage AI technologies responsibly and equitably in language teaching and learning.

Cite this article: Kianinezhad, N. (2025). Review of AI in Language Teaching, Learning, and Assessment (2024) edited by Fang Pan. IGI Global: 384 pp., ISBN: 9798369308738. *Applied Linguistics Inquiry*, *3*(1), 162-164. doi: 10.22077/ali.2025.8602.1059

2. Structural organization

2.1. Preview: Foundations and context

The first section, Preview sets the foundation for understanding the transformative potential of AI in language education. Pan explores the paradigm shift from traditional, teacher-centered methods to innovative, AI-enabled pedagogical models. This shift includes tools such as interactive chatbots, adaptive learning systems, and automated assessments that personalize learning by tailoring instruction to the unique needs of each student.

Pan further underscores the importance of collaboration among educators, AI developers, and policymakers in ensuring the seamless and ethical integration of these technologies. He advocates for sustained empirical research as a key driver in assessing AI's effectiveness, addressing its limitations, and solidifying its long-term role in modern education. By building this strong conceptual foundation, Pan prepares readers to engage with AI's more intricate applications and challenges in subsequent sections.

2.2. Perspectives: Diverse and global insights

The Perspectives section provides a critical exploration of AI from a variety of cultural, educational, and policy-related standpoints. Drawing on insightful case studies from regions such as Malaysia and Hong Kong, Pan effectively highlights the dual nature of AI integration—the opportunities it brings, as well as the challenges it poses.

For instance, learners in Malaysia applaud AI for its ability to deliver personalized, student-centered learning experiences, which were previously unattainable due to resource constraints. Conversely, educators in Hong Kong express concerns over the potential displacement of human teachers, raising essential questions about AI's impact on teacher-student dynamics and pedagogical relationships. By weaving these global perspectives together, Pan presents a nuanced analysis of the multifaceted challenges and opportunities that accompany AI implementation across diverse educational systems.

2.3. Practice: Practical applications and innovations

In the Practice section, Pan shifts focus toward AI's tangible contributions to language education, showcasing concrete, real-world applications through a series of compelling case studies. Key innovations include adaptive language tutoring systems, which offer real-time personalized feedback tailored to individual learner needs, and intercultural communication simulations, which immerse students in AI-driven environments to develop practical intercultural competencies.

Additionally, the book explores AI-enhanced content development, where machine learning tools produce dynamic and adaptive teaching materials, ensuring more effective and inclusive instruction. For younger learners, Pan highlights the transformative role of AI-driven storytelling tools that engage children through interactive and creative narratives, fostering both motivation and participation. Furthermore, AI is shown to align language learning curricula with the United Nations Sustainable Development Goals (SDGs), promoting inclusive global education and nurturing global citizenship.

By addressing challenges such as scalability, resource limitations, and student motivation, the Practice section compellingly demonstrates AI's potential to enhance accessibility and optimize learning outcomes in diverse settings.

3. Key themes

3.1. Personalized learning

A dominant theme throughout the book is AI's capacity to revolutionize education through personalized learning. By leveraging adaptive systems, AI effectively tracks learner progress, identifies individual needs, and tailors instructional content accordingly. This student-centered approach not only enhances learning outcomes but also significantly boosts learner motivation and engagement, ensuring an inclusive educational experience.

3.2. Ethical consideration

Pan also pays considerable attention to the ethical dimensions of AI adoption. The book highlights concerns surrounding data privacy, algorithmic bias, and the transparency of AI-driven systems. While AI presents exciting possibilities, Pan underscores the need for ethical accountability and inclusive policies to prevent exploitation or inequities. Accordingly, the book calls for collaborative action among educators, policymakers, and developers to establish robust frameworks that prioritize fairness and accessibility.

3.3. Instruments

AI emerges as a transformative force capable of redefining traditional pedagogical approaches. Tools such as virtual teaching assistants and gamified platforms foster highly interactive and student-centered classrooms that cater to diverse learning preferences. Nevertheless, Pan acknowledges persistent barriers, including digital literacy gaps, implementation costs, and

resistance to technological change. As such, he emphasizes the importance of comprehensive training and institutional support to ensure educators are equipped to navigate AI adoption effectively.

3.4. Strength of the volume

Pan's volume possesses several distinctive strengths that make it invaluable to its target audience. First, it's clear and logical structure progressing from foundational theories to applied innovation ensures clarity and facilitates deep understanding. Additionally, the interdisciplinary approach, drawing insights from educators, researchers, and technologists, provides a well-rounded perspective on AI integration. Finally, the inclusion of practical, real-world case studies enhances the book's relevance by demonstrating actionable solutions and inspiring innovative approaches to educational challenges.

4. Potential limitations

Despite its many strengths, the book could benefit from further exploration in two specific areas. Firstly, a greater emphasis on AI literacy for stakeholders, particularly educators and students, would be valuable in preparing them to adopt and adapt to AI tools effectively. Secondly, there is a need for a more transdisciplinary approach—one that begins with a central theme or issue and then draws on related disciplines to address it comprehensively. This approach could enrich the exploration of AI's role in language education, integrating a variety of perspectives to better tackle complex challenges. Additionally, a deeper discussion regarding socio-economic and cultural barriers, particularly in low-resource educational settings, would help ensure that AI's benefits are distributed equitably and inclusively across global communities.

As a final note, AI in Language Teaching, Learning, and Assessment is a meticulously researched, well-structured, and deeply insightful volume that addresses the profound impact of AI on language education. By balancing optimism for AI's transformative potential with a thoughtful exploration of its ethical and practical complexities, Pan offers readers a timely roadmap for AI integration.

This book skillfully combines theoretical frameworks, diverse global perspectives, and practical applications to empower educators, researchers, and policymakers alike. It serves as both a strategic guide and a call to action, encouraging stakeholders to harness AI's vast possibilities while upholding human oversight, transparency, and equity. In fact, for anyone invested in the future of language education, this work stands as an invaluable contribution that illuminates the path forward.

5. References

Kianinezhad, N. (2023). Teacher development in technology-enhanced language teaching: Book review. *Journal of Language and Education*, *9*(4), 180-184. https://doi.org/10.17323/jle.2023.17676

Pan, F. (Ed.). (2024). *AI in Language Teaching, Learning, and Assessment*. IGI Global. https://doi.org/10.4018/979-8-3693-0872-1