



JOURNAL OF SECOND LANGUAGE PEDAGOGY

JOURNAL OF SECOND LANGUAGE PEDAGOGY

JSLP Volume1 , Issue 3

2025



Table of Contents

1. Roles of Aural and Orthographic Vocabulary Knowledge in Intermediate EFL Learners' Listening Comprehension	
Bitra Lamouchi Tarazi, Mohammad Taghi Farvardin*.....	1-12
2. Language Teachers' Professional Identity Related Studies: A Narrative Review	
Matin Torkchin *, Mohammad Reza Baradari.....	13-26
3. Secondary- level EFL Teachers' Perceptions toward Doing Research and Their Challenges in Research Engagement	
Sara Noroozi, Effat Fathi, Gholamreza Zareian*, Mohammad Reza Amirian.....	27-45
4. The Effect of ChatGPT Implementation on Right- and Left-Hemisphere Dominant Iranian EFL Learners' Argumentative Writing and Metacognitive Awareness	
Hassan Alizadeh Mahmoud Alilo*.....	46-66
5. A Narrative Review of Textbook Checklists: Insights for Language Pedagogy	
Hamidreza Zahedi Gholezo*, Maryam Ghadamyari.....	67-82
6. A Feasibility Study of Lesson Study Implementation for Iranian English Teachers' Professional Development: Managers' Insights	
Reza Hadinezhad, Seyed Mohammad Reza Amirian*, Mohammad Davoudi.....	83-100



Contents lists available at [JSLP](https://www.sanad.iau.ir/journal/jslp)

Journal of Second Language Pedagogy

Journal homepage: <https://www.sanad.iau.ir/journal/jslp>

Roles of Aural and Orthographic Vocabulary Knowledge in Intermediate EFL Learners' Listening Comprehension

Bitā Lamouchi Tarāzi¹, Mohammad Taghi Farvardin^{2*}

¹ Department of English, Ahv.C., Islamic Azad University, Ahvaz, Iran

² Department of English, Ahv.C., Islamic Azad University, Ahvaz, Iran

*Corresponding Author's Email: mtfarvardin@iau.ac.ir

KEY TERMS

Aural vocabulary knowledge
Intermediate EFL learners
Listening comprehension
Orthographic vocabulary knowledge

ABSTRACT

Although vocabulary knowledge is essential for learning a foreign language, the distinct contribution of aural and orthographic vocabulary knowledge to listening comprehension has not received enough attention. This study investigated the roles of aural and orthographic vocabulary knowledge in intermediate English as a foreign language (EFL) learners' listening comprehension. To this end, 112 intermediate EFL learners were selected through convenience sampling, using the Oxford Placement Test to determine their proficiency level. The tests used in this study were the Vocabulary Levels Test (VLT), the Aural Vocabulary Knowledge (AVK) test, and a listening comprehension test. The scores of the AVK test and those of the listening comprehension test revealed a strong correlation ($r = .63$, $p < .01$); the VLT scores showed a moderate correlation with listening comprehension scores ($r = .41$, $p < .01$). Regression analysis further revealed that AVK was a more effective predictor of English listening comprehension ($\beta = .48$, $p < .01$). In other words, the listening comprehension of intermediate EFL learners depended more on AVK than on orthographic vocabulary knowledge. The results suggest that increasing EFL learners' listening comprehension mostly depends on AVK, which can have a key role in their communicative competence.

ARTICLE TYPE

Original Research Paper

Received:	8 May 2025
Revised:	20 June 2025
Accepted:	25 June 2025
Published Online:	1 July 2025

© The Authors 2025

1. Introduction

Listening comprehension as a basic component of English as a Foreign Language (EFL) proficiency, is a tool actively used by students in social, academic, and professional settings (Vandergriff & Goh, 2012). Its importance stems from its main role as a main facilitator of interactive communication. EFL learners frequently struggle with comprehending conversations and talks by native speakers (Dunkel, 1991; Field, 2008). Often these challenges relate to different aspects of vocabulary knowledge, that is, both the word number a person knows and how well they grasp those words (Clenton et al., 2025;

Nation, 2001). Although most agree that vocabulary size is basic for language development, the special contributions of aural (spoken) and orthographic (written) vocabulary knowledge to listening skills have remained understudied, especially in settings where instructional practices emphasize the written forms (Matthews & Lange, 2024). This highlights the need of knowing how aural and written vocabulary knowledge can each help EFL learners, especially those at the lower levels of language proficiency. The multi-dimensionality of lexical skill complicates this task. Effective listening not only demands semantic understanding but also rapid recognition of phonological and orthographic forms, which together enable learners to decode spoken input in real time (Goh, 2000, 2002; Perfetti, 2007; Schmitt, 2010).

Learners with strong aural vocabulary knowledge, for example, or those who can identify and recall word meaning from listening, are likely to be more capable of processing spoken language, hence lowering cognitive load and freeing mental resources for higher-level operations including inferencing or building contextual cues (Vandergrift & Goh, 2012). Conversely, overreliance on orthographic skills, often fostered in education systems that place strong emphasis on written exams, can hinder listening fluency. This is because written forms do not directly relate to the ephemeral, often contracted nature of speech, which offers features such as contractions, elision, and suprasegmental ones like intonation and stress (Ghorbani Nejad & Farvardin, 2022; Matthews & Lange, 2024). In settings like Iran, where grammar-translation approaches govern classrooms, students suffer from phonological blindness—a term Rost (2011) has developed to explain why students attribute listening difficulty to unfamiliar vocabulary rather than unseen connected speech routines. This creates an imbalance of lexis whereby orthographic knowledge overrides phonological knowledge. The interdependence between vocabulary knowledge and listening comprehension has given rise to many debates among scholars (e.g., Matthews & Cheng, 2015; Vandergrift, 2006, 2007).

Neurological awareness science, such as research based on Baddeley's (1983) theory of working memory, shows that the phonemic cycle is an essential subsystem to process listening information. This memory system retains and broadcasts information orally in short-term memory, allowing listeners to link listening signals with internal vocabulary. Particularly in the acquisition of a second language (L2), aural vocabulary knowledge is essential for effective listening comprehension and general language proficiency (Matthews & Cheng, 2015; Matthews & Lange, 2024). Since most daily communication takes place in spoken interaction, learners' success in using real-life language depends directly on their capacity to identify and comprehend vocabulary aurally (Matthews & Lange, 2024). Studies have indicated that a learner's aural vocabulary (Vandergrift & Goh, 2012) greatly influences listening comprehension; insufficient aural lexical knowledge can thus impede processing spoken input in real time. Furthermore, aural vocabulary provides the basis for fluency development, which helps students to decode, interpret, and react correctly in communicative environments (Milton & Fitzpatrick, 2013). Therefore, aural vocabulary knowledge is quite important in social and academic spheres of language use, which emphasizes the need for tests aimed at this particular ability (Goh, 2000, 2002; Harsch, 2014; Milton et al., 2010).

EFL learners at lower levels of proficiency find great challenges in such settings. Elementary and intermediate learners struggle with partial vocabulary knowledge, while advanced learners can use metacognitive strategies for overcoming phonemic knowledge (Elgort et al., 2018). Their limited vocabulary affects real-time comprehension; hence, instead of comprehending long words, they must use too much mental effort in decoding individual words (Matthews & Cheng, 2015; Matthews & Lange, 2024; Wang & MacIntyre, 2021). For the learners who are rarely exposed to lowered phonemes and

the high speaking rate of normal communication, the issue becomes more demanding, especially in the settings where contact with English speakers is limited.

Uchihara et al. (2019) examined the predictive value of aural and orthographic word knowledge for 150 Japanese EFL learners. They used regression analyses and found that knowledge of aural vocabulary ($\beta = .52$) was a stronger predictor of listening comprehension than orthographic knowledge ($\beta = .24$) even when overall proficiency was accounted for. This study attributed these results to the straightforward mapping of phonological processing to listening demands, which corroborated the present findings. They also identified orthographic knowledge indirectly enhancing comprehension through reading-based inferencing, a finding suggesting modality-specific synergies at advanced levels of proficiency. In China, Qiu and Luo's (2022) quasi-experimental study of flipped listening instruction examined pre-class video modules with subsequent in-class interactive tasks appreciably improving listening performance and reducing anxiety in intermediate learners. The emphasis of the flipped method on independent learning allowed students to go over challenging phonological elements, thereby strengthening auditory vocabulary. However, the study's small sample size ($N=44$) and short intervention duration (a single semester) placed a limit on generalizability, calling into question the scalability of such practice in low-resource settings. Similarly, He et al.'s (2022) study on a sample of 797 Chinese students revealed radical modality-specific differences: students performed well on reading comprehension but poorly on listening, which was accounted for by limited practice in listening input and inadequate phonological training. The findings reinforce Stæhr's (2009) earlier study that found orthographic vocabulary moderately related to listening accomplishment in higher-level learners but not intermediate, suggesting written knowledge is beneficial based on proficiency level.

In Iran, the preeminence of standardized tests—e.g., the University Entrance Exam, which gives high priority to grammar and written vocabulary—shapes what is taught, often at the expense of the development of aural skills (Rezai, 2023). Teachers, constrained by rigid curricula and little training in communicative methods, fall back on grammar-translation methods that prioritize translation and memorization over interactive practice listening. Such an approach creates a self-reinforcing cycle where students perform well on paper tests but lag behind in real listening situations, perpetuating the stereotypical belief that listening is a lower priority than reading or writing.

Elgort et al. (2018) explored orthographic processing's role in listening comprehension across levels of proficiency. Their experiment revealed that strong orthographic knowledge hindered novices' phonological decoding by strengthening written form dependency, yet aided advanced learners through subvocalization processes. Through the examination of the following research questions, this study aimed to address the existing gap in the literature:

RQ1: To what extent do intermediate EFL learners' aural vocabulary knowledge test scores and listening comprehension test scores correlate?

RQ2: To what extent do intermediate EFL learners' orthographic vocabulary knowledge test scores and listening comprehension test scores correlate?

RQ3: To what extent do aural and orthographic vocabulary knowledge predict intermediate EFL learners' listening comprehension?

2. Methodology

This study adopted a quantitative, correlational design. This section includes the description of the participants, the instruments, and the data collection procedure.

2.1 Participants and Setting

In this study, 112 intermediate EFL learners aged 16–25 years old from three private language institutes in Ahvaz, Iran, were recruited. Participants were selected through convenience sampling. To ensure homogeneity of proficiency level, all participants took part in the Oxford Placement Test (OPT), and only those in the intermediate band were chosen (i.e., 112 were selected out of 130 EFL learners). Participants were of an evenly balanced demographic profile of high school students (52%), university undergraduates (48%). Gender balance was relatively equal, at 58 female and 54 male EFL learners, to ensure that outcomes were not inappropriately influenced by gender.

2.2 Instrumentation

Participants' aural vocabulary knowledge (AVK) was assessed using a test adapted from Matthews (2018). The test included three levels of word frequency: 23 words from level one (0–2000 frequency range), 27 words from level two (2001–3000 frequency range), and 13 words from level three (3001–5000 frequency range). Its purpose was to evaluate the test takers' competency in both comprehension and production of a total of 63 target words. Words from level one and level two were categorized as highly frequent, while level three words were classified as mid-frequency (Schmitt & Schmitt, 2014). Participants, after listening to a stimulus sentence, were required to select one target word for each item on the AVK test. The test paper included a contextual sentence for each item with a blank space designated for the corresponding target word in written form. Test-takers listened to each sentence only once.

Orthographic vocabulary was assessed using Nation's Vocabulary Levels Test (VLT) (1990), a reliable written test where participants matched 60 target words (like benefit and obtain) to their meanings at 2000-word and 3000-word frequency levels. For example, benefit had choices like advantage or obstacle preceding it with written form—meaning association without context cues. The stratified design of VLT enabled lexical proficiency to be examined at differing levels of proficiency, and pilot adjustments enabled straightforward definitions for intermediate-level learners that were highly reliable ($\alpha = .82$).

Third, listening comprehension was assessed using a test that included six real audio recordings, like lectures and customer service calls, followed by 24 multiple-choice questions that checked for understanding, remembering details, and identifying the main idea. All tests were piloted with 20 students to determine item clarity and timing, which resulted in Cronbach's alpha values above .84, which confirms the reliability.

2.3 Research Procedure

Data collection followed a systematic, multi-phase procedure to ensure rigor and minimize bias. Participants first completed the OPT to establish intermediate proficiency. Only learners who achieved scores in the intermediate band progressed to the main study. Eligible participants

subsequently completed the three tests across three sessions, with the order of the aural and orthographic tests randomized to counteract order effects. For instance, half of the participants did the aural test before the other half did the orthographic test. Both tests were done in a quiet classroom environment to minimize distractions, with explicit verbal and written instructions. Breaks of 5–10 minutes were enforced between sections to counteract fatigue. Ethical practice was rigorously followed: participants provided written informed consent, data were anonymized using ID codes (e.g., P01, P02), and results were stored confidentially for secure storage.

3. Data Analysis

In this section, the results of the normality tests, descriptive statistics, and correlational analyses are presented. First, the assumptions of normality were assessed using the Shapiro-Wilk test. Results indicated no significant deviations from normality for any variables ($p > .05$), supporting the use of parametric analyses (see Table 1).

Table 1

Results of Shapiro-Wilk Normality Test

Variable	W	p
Aural Vocabulary	0.38	.19
Orthographic Vocabulary	0.26	.12
Listening Comprehension	0.21	.09

Table 1 presents the results of the Shapiro-Wilk test, which established whether the distributions of the three principal variables (aural vocabulary, orthographic vocabulary, and listening comprehension) significantly deviated from normal. The W statistic and respective p-values are given. Since all p-values were greater than the .05 significance level, the null hypothesis of normality was retained, confirming that parametric tests were appropriate for subsequent analysis. Table 2 presents means and standard deviations of the three test scores.

Table 2

Results of Descriptive Statistics

Variable	M	SD
Aural Vocabulary (63)	31.5	5.3
Orthographic Vocabulary (60)	38.2	6.5
Listening Comprehension (24)	15.7	3.4

Note. Maximum possible scores are indicated in parentheses.

Table 2 depicts the central tendency and variability of scores for aural vocabulary, orthographic vocabulary, and listening comprehension. The maximum possible scores for each test are provided in parentheses to contextualize performance. For example, the aural vocabulary test had a maximum score of 63, with participants scoring an average of 31.5 ($SD = 5.3$). The mean score of the VLT was ($M = 38.2$, $SD = 6.5$). The participants' mean score in the listening comprehension test was 15.7 ($SD = 3.4$). Pearson correlations revealed significant relationships between the variables (Table 3).

Table 3

Results of Pearson Correlations

Variable	1	2	3
1. Aural Vocabulary	—		
2. Orthographic Vocabulary	.38**	—	
3. Listening Comprehension	.63**	.41**	—

Note. ** $p < .01$

Table 3 displays the correlations between the three variables. Aural vocabulary knowledge showed a moderate positive correlation with orthographic vocabulary ($r = .38$, $p < .01$) and a strong and positive correlation with listening comprehension ($r = .63$, $p < .01$). Orthographic vocabulary also correlated moderately with listening comprehension ($r = .41$, $p < .01$). A multiple linear regression examined the predictive power of aural and orthographic vocabulary on listening comprehension (Table 4).

Table 4

Results of Regression Analysis for Listening Comprehension Prediction

Predictor	B	SE B	β	t	p	95% CI
Aural Vocabulary	0.48	0.09	.48	5.33	<.01	[0.30, 0.66]
Orthographic Vocabulary	0.12	0.07	.12	1.71	.18	[-0.02, 0.26]

Model Summary: $R^2 = .52$

Table 4 details the results of the multiple linear regression analysis predicting listening comprehension scores from aural and orthographic vocabulary knowledge. The unstandardized (B) and standardized (β) coefficients, standard errors (SE B), t-values, p-values, and 95% confidence intervals (CI) are reported. Aural vocabulary emerged as a significant predictor ($\beta = .48$, $p < .01$), with each unit increase in aural scores associated with a 0.48 standard deviation increase in listening performance. Orthographic vocabulary was not a significant predictor ($\beta = .12$, $p = .18$). The model explained 52% of the variance in listening scores ($R^2 = .52$), as indicated by the significant F-statistic ($F(2, 109) = 32.14$, $p < .01$).

4. Discussion and Conclusion

The results showed the key role of aural vocabulary knowledge in intermediate EFL learners' listening comprehension. The strong correlation between aural vocabulary and listening comprehension ($r = 0.63$) is consistent with the hypothesis of Perfetti's lexical quality (2007), showing high-quality phonemic knowledge is essential for listening comprehension. Students with excellent audio vocabulary can decode more effectively, reduce perception efforts, and release more attention for more order-like reasoning or context integration. These results are particularly relevant to educational systems such as Iran, where traditional education emphasizes accuracy in writing related to mastery of the mouth, creating a gap between daily learning and communication requirements. For example, Iranian students work well in writing tests but very badly in the authentic listening tasks because their education emphasizes memorizing spelling forms rather than exposure to spoken language. This system leads to what Rost (2011) described as phonetic blindness, meaning students struggle with understanding spoken language and instead focus on analyzing unfamiliar words.

The results of this study confirm the significant role that aural vocabulary knowledge plays in helping intermediate EFL learners to support their listening comprehension. The strong correlation noted between AVK scores and listening performance ($r = .63$) is in line with previous study by Uchihara et al. (2019), who also found aural vocabulary to be the stronger predictor of listening comprehension compared to orthographic knowledge. Given real-time spoken input requiring quick decoding of acoustic signals, the results support the claim made by Vandergrift and Goh (2012) that phonological processing is more directly linked to listening skill.

The findings also confirm the strong influence of AVK using regression analysis ($\beta = .48$). The use of intermediate learners closes the gap in a literature dominated by advanced learners. Unlike advanced learners who use metacognitive strategies like predicting content or monitoring themselves to compensate for auditory gaps (Vandergrift & Goh, 2012), intermediate learners struggle with phonological representation of words. This difference is explained by Segalowitz's (2005) distinction between knowledge—static word recognition—and skill—dynamic processing under time pressure. An intermediate learner might, for example, correctly define a word in writing but fail to identify it in a rapid academic lecture, so illustrating the difference between declarative knowledge and procedural skill. Emphasizing repeated aural exposure and targeted phonological drills, the situation highlights the need for pedagogical methods that move students from memorization and toward automatic recall. Learners in Iran, where grammar-translation techniques predominate in classroom instruction, receive no practice in parsing connected speech and therefore are not ready for the reduced forms and suprasegmental features of natural spoken English.

By contrast, knowledge of orthographic vocabulary showed a lower but still statistically significant correlation with listening ($r = .41$). This reflects the results of Staehr (2009) and Elgort et al. (2018), who found, especially at lower and intermediate proficiency levels, written vocabulary has only moderate correlation with listening comprehension. An overreliance on written forms can, as Elgort et al. (2018) observed, impede phonological decoding at early stages by reinforcing visual word recognition strategies inappropriate for real-time auditory processing. The lower contribution of orthographic knowledge in the current study most likely results from the instructional bias in Iran, where grammar-translation and exam-oriented approaches limit students' exposure to natural

speech patterns and lower the functional development of listening-related skills (Ghorbani Nejad & Farvardin, 2022).

The cognitive and neurological foundations of language processing offer still another justification for these findings. Baddeley's (1983) working memory model holds that good listening calls for active use of the phonological loop to momentarily store and rehearse spoken input. Stronger aural vocabulary knowledge is probably more effective in this mental process, which can help learners decode, segment, and understand speech with less cognitive load. This can help the lower-level EFL learners, who are in a developmental stage and phonological decoding skills are still developing.

This study backs the recommendations of Matthews and Lange (2024) that enhancing auditory lexical access enhances working memory capacity for higher-order comprehension activities, including contextual prediction and inferencing. These results also support the mounting corpus of data that is pushing pedagogical change in EFL environments. Instructional strategies need to be changed to give spoken language exposure top priority since listening comprehension depends more on aural than orthographic vocabulary knowledge. Research, including Qiu and Luo (2022) and He et al. (2022), has shown that direct listening instruction—supported by interactive and phonologically focused strategies—yields better comprehension outcomes than conventional written-centric models.

Orthographic vocabulary knowledge tends to support skills like reading, but it may also contribute indirectly to listening comprehension. Learners who have well-developed orthographic vocabulary might use this knowledge to engage in mental rehearsal or prediction when listening, especially in academic or test-based settings where prior reading of related texts is common (Clenton et al., 2025). Still, such transfer is often partial and less immediate than aural access. Studies such as Uchihara et al. (2019) have demonstrated that aural vocabulary knowledge has a stronger predictive value for listening comprehension than orthographic vocabulary, particularly among intermediate learners. This is likely because the phonological nature of listening demands rapid auditory decoding that written knowledge alone cannot sufficiently support.

Furthermore, a mismatch between the two types of vocabulary knowledge can lead to processing difficulties (Ghorbani Nejad & Farvardin, 2022; Matthews & Lange, 2024). Learners may recognize a word in writing but not in speech, especially if the spoken form is unfamiliar due to pronunciation differences or lack of phonological training. This discrepancy is especially pronounced in EFL settings, where learners are frequently exposed to words in their written form but rarely hear them pronounced by native or fluent speakers (Goh, 2000, 2002; Harsch, 2014; Milton et al., 2010). In addition, the lexical quality hypothesis (Perfetti, 2007) suggests that high-quality lexical representations include well-integrated form, meaning, and use. Moreover, in line with the findings of this study, it can be argued that aural vocabulary knowledge contributes directly to the phonological aspect of this representation, enhancing the learner's ability to integrate spoken words into meaningful interpretations during listening (Matthews & Lange, 2024). When lexical items are automatically recognized, listeners can allocate more attention to higher-order comprehension processes such as inference-making and interpreting speaker intent (Vandergrift & Goh, 2012). In contrast, poor aural vocabulary knowledge forces learners to focus excessively on decoding, which detracts from overall comprehension.

The frequency of word exposure also plays a critical role in vocabulary development and listening proficiency. High-frequency words are accessed more quickly and reliably, particularly in spoken form, and are essential for achieving fluency (Uchihara, 2019). Instructional environments that emphasize written materials may offer frequent exposure to orthographic forms but insufficient practice with their aural counterparts. Another theoretical construct relevant here is incidental vocabulary learning through listening. When learners engage with spoken input that is slightly above their current proficiency level, they acquire new vocabulary through context-driven inference and repeated exposure (Clenton et al., 2025). This process is more effective when learners have a foundational aural vocabulary that enables them to follow the discourse and isolate new lexical items.

There is also growing evidence that the modality in which vocabulary is learned affects its retrieval. Words learned through listening are more readily recalled in listening tasks, while those learned through reading are more accessible during reading (Matthews, 2018; Matthews & Cheng, 2015; Matthews & Lange, 2024). This modality-specificity supports the need to balance both forms of vocabulary instruction to ensure comprehensive proficiency. In contexts dominated by grammar-translation approaches, learners often have highly developed orthographic vocabulary knowledge but underdeveloped phonological knowledge. This imbalance can lead to difficulties in understanding connected speech, idiomatic expressions, and reduced forms—all of which are essential elements of fluent listening comprehension (Rost, 2011).

The Iranian setting, which is dominated by test-driven methods and little real listening, emphasizes how urgently this change is needed. Teachers have to use more balanced strategies that foster both listening input and vocabulary development in their auditory form—an approach indispensable for closing the distance between classroom learning and actual communication. Furthermore, taken into account is the fact that the intermediate proficiency level of the participants could have influenced the relative significance of aural versus orthographic vocabulary. While lower-level learners depend more on auditory input due to limited decoding skills, advanced learners can use orthographic knowledge strategically through subvocalization and mental rehearsal (Elgort et al., 2018).

The superiority of aural vocabulary knowledge in predicting listening comprehension suggests that learners not yet developing the cognitive strategies or lexical depth need to use written vocabulary as a successful listening aid. Beginning with intensive aural training and progressively including orthographic reinforcement, as learners' proficiency grows, can support instructional strategies (Clenton et al., 2025). The strong predictive ability of aural vocabulary knowledge revealed in this study emphasizes the need to review conventional EFL courses that sometimes exclude listening instruction in favor of reading and grammar.

Learners submerged in systems that favor orthographic input create a distorted lexical profile, which fails to meet the demands of real-time spoken communication, as Matthews and Lange (2024) argued. This mismatch between instruction and practical language use not only reduces learners' communicative competence but also maintains an underperformance cycle in listening assignments. The present results support a more balanced, input-rich pedagogical approach that emphasizes phonological form recognition, integrates frequent exposure to real spoken English, and promotes vocabulary development using both bottom-up and top-down listening strategies. Including listening-oriented tests and assignments fit for the aural lexical demands of daily English can help

teachers close the gap between classroom instruction and communicative efficacy, enabling more autonomous and competent language users.

The strong correlation between listening capacity and aural vocabulary ($R = 0.63$) supports the highlight of the treatment of phonemes in real-time understanding, strengthening theoretical accounts such as the hypothesis of lexical quality (Perfetti, 2007) and Baddeley's (1983) working memory model. These results indicate that the ability of learners to identify and restore word forms reduces the cognitive load and prepares mental resources for high-level processes such as inference. On the other hand, orthographic vocabulary was found to be moderately correlated with listening comprehension. This gap is particularly clear in systems focusing on tests, such as Iran's, where students write good but poor tests in actual listening activities due to contact with the aspects of the connected words, such as shrinking and ultrasound.

Phonological awareness, such as listening tasks using natural speech documents including podcasts, TED Talks, or films, should be prioritized by teachers. Distinguishing tasks of phonemes—for example, the difference between minimal pairs such as ship and sheep or identifying syllables emphasized in polysyllabic words—can refine the decoding ability of learners. Spelling exercises, where learners read less than what is understood, can also facilitate the mapping of forms in a sensitive situation over time, simulating future listening requirements. These methods are linked to the involvement load hypothesis (Laufer & Hulstijn, 2001), which suggests that how a task can help store vocabulary more effectively. However, the implementation of such reforms in systems based on exams, such as Iran, requires a big change. Program designers can reduce excessive writing tests by combining reviews based on listening to real scenarios, such as summarizing audio conferences or answers to loud questions. Teachers' training programs should provide teachers with interactive listening education, including training on technology use, such as language programs with instant feedback options, or to create multimedia lessons combined with images and audio. Intermediate learners may receive these reforms to fill the gap between learning and the real application of language, turning passive learning into more practical one.

The results imply that expanding L2 listening comprehension for students mostly depends on AVK as a main construct of vocabulary knowledge. Supporting other facets of the language and greatly improving learners' L2 listening comprehension seem to depend on a strong lexical basis. Listening comprehension is knowing not only words but also context, idioms, and details. Consequently, increasing a strong vocabulary by diverse exposure and practice can result in more efficient listening comprehension, so enhancing the general communicative competence.

Despite these findings, this research has its own limitations to be taken into account. Replicating studies in different EFL contexts can enhance the research generalization. Research on emotional variables, such as anxiety or motivation, can also reduce their interaction with vocabulary and listening performance. Moreover, examining the roles of different types of vocabulary knowledge across proficiency levels can be insightful and therefore, recommended. Finally, conducting mixed-methods studies are recommended to meticulously explore the roles of aural and orthographic vocabulary knowledge in EFL learners' listening comprehension.

References

- Baddeley, A. D. (1983). Working memory. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 302(1110), 311-324. <https://doi.org/10.1098/rstb.1983.0057>
- Clenton, J., Brooks, G., & Laufer, B. (2025). Vocabulary studies in L1 and L2 development: The interface between theory and practice. *Languages*, 10(4), 81-94. <https://doi.org/10.3390/languages10040081>
- Dunkel, P. (1991). Listening in the native and second/foreign language: Toward an integration of research and practice. *TESOL Quarterly*, 25(3), 431-457. <https://doi.org/10.2307/3586979>
- Elgort, I., Brysbaert, M., Stevens, M., & Van Assche, E. (2018). Contextual word learning during reading in a second language: An eye-movement study. *Studies in Second Language Acquisition*, 40(2), 341-366. <https://doi.org/10.1017/S0272263117000109>
- Field, J. (2008). *Listening in the language classroom*. Cambridge University Press.
- Ghorbani Nejad, S., & Farvardin, M. T. (2022). Roles of general language proficiency, aural vocabulary knowledge, and metacognitive awareness in L2 learners' listening comprehension. *International Journal of Listening*, 36(3), 187-206. <https://doi.org/10.1080/10904018.2019.1572510>
- Goh, C. (2000). A cognitive perspective on language learners' listening comprehension problems. *System*, 28(1), 55-75. [https://doi.org/10.1016/S0346-251X\(99\)00060-3](https://doi.org/10.1016/S0346-251X(99)00060-3)
- Goh, C. C. M. (2002). Exploring listening comprehension tactics and their interaction patterns. *System*, 30(2), 185-206. [https://doi.org/10.1016/S0346-251X\(02\)00004-0](https://doi.org/10.1016/S0346-251X(02)00004-0)
- Harsch, C. (2014). General language proficiency revisited: Current and future issues. *Language Assessment Quarterly*, 11(2), 152-169. <https://doi.org/10.1080/15434303.2014.902059>
- He, L., Xiong, L., & Min, S. (2022). Diagnosing listening and reading skills in the Chinese EFL context: Performance stability and variability across modalities and performance levels. *System*, 106, 102-110. <https://doi.org/10.1016/j.system.2022.102787>
- Laufer, B., & Hulstijn, J. (2001). Incidental vocabulary acquisition in a second language: The construct of task-induced involvement. *Applied Linguistics*, 22(1), 1-26. <https://doi.org/10.1093/applin/22.1.1>
- Matthews, J. (2018). Vocabulary for listening: Emerging evidence for high and mid-frequency vocabulary knowledge. *System*, 72, 23-36. <https://doi.org/10.1016/j.system.2017.10.005>
- Matthews, J., & Cheng, J. (2015). Recognition of high frequency words from speech as a predictor of L2 listening comprehension. *System*, 52, 1-13. <https://doi.org/10.1016/j.system.2015.04.015>
- Matthews, J., & Lange, K. (2024). Inadequate aural vocabulary knowledge of high-frequency words. *CALICO Journal*, 41(2), 99-121. <https://doi.org/10.1558/cj.23455>
- Milton, J., & Fitzpatrick, T. (2013). *Dimensions of vocabulary knowledge*. Palgrave Macmillan.

- Milton, J., Wade, J., & Hopkins, N. (2010). Aural word recognition and oral competence in a foreign language. In R. Chacón-Beltrán, C. Abello-Contesse, & M. Torreblanca-López (Eds.), *Insights into non-native vocabulary teaching and learning* (pp. 83–98). Bristol, UK: Multilingual Matters.
- Nation, I. S. P. (1990). *Teaching and learning vocabulary*. Heinle & Heinle.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Qiu, Y., & Luo, W. (2022). Investigation of the effect of flipped listening instruction on the listening performance and listening anxiety of Chinese EFL students. *Frontiers in Psychology*, 13, 104–130. <https://doi.org/10.3389/fpsyg.2022.1043004>
- Rezai, A. (2023). Online peer-dynamic assessment as an approach for promoting listening comprehension and relieving learning anxiety: Evidence from Iranian EFL learners. *Computer-Assisted Language Learning Electronic Journal*, 24(2), 6–31. <https://callej.org/index.php/journal/article/view/5>
- Rost, M. (2011). *Teaching and researching listening* (2nd ed.). Pearson Education.
- Schmitt, N. (2010). *Researching vocabulary: A vocabulary research manual*. Palgrave Macmillan.
- Schmitt, N., & Schmitt, D. (2014). A reassessment of frequency and vocabulary size in L2 vocabulary teaching. *Language Teaching*, 47(4), 484–503. <https://doi.org/10.1017/S0261444812000018>
- Segalowitz, N. (2005). Automaticity and second language learning. In J. F. Kroll & A. M. B. De Groot (Eds.), *Handbook of bilingualism: Psycholinguistic approaches* (pp. 371–388). Oxford University Press.
- Stæhr, L. S. (2009). Vocabulary knowledge and advanced listening comprehension in English as a foreign language. *Studies in Second Language Acquisition*, 31(4), 577–607. <https://doi.org/10.1017/S0272263109990039>
- Uchihara, T., Webb, S., & Yanagisawa, A. (2019). The effects of repetition on incidental vocabulary learning: A meta-analysis of correlational studies. *Language Learning*, 69(3), 559–599. <https://doi.org/10.1111/lang.12343>
- Vandergrift, L. (2006). Second language listening: Listening ability or language proficiency? *The Modern Language Journal*, 90(1), 6–18. <https://doi.org/10.1111/j.1540-4781.2006.00381.x>
- Vandergrift, L. (2007). Recent developments in second and foreign language listening comprehension research. *Language Teaching*, 40(3), 191–210. <https://doi.org/10.1017/S0261444807004338>
- Vandergrift, L., & Goh, C. (2012). *Teaching and learning second language listening: Metacognition in action*. Routledge.
- Wang, L., & MacIntyre, P. D. (2021). Second language listening comprehension: The role of anxiety and enjoyment in listening metacognitive awareness. *Studies in Second Language Learning and Teaching*, 11(4), 491–515. <https://doi.org/10.14746/ssllt.2021.11.4.2>



Contents lists available at [JSLP](#)

Journal of Second Language Pedagogy

Journal homepage: <https://www.sanad.iau.ir/journal/jslp>

Language Teachers' Professional Identity Related Studies: A Narrative Review

Matin Torkchin^{1*}, Mohammad Reza Baradari²

¹ Department of English and Literature, Allameh Tabataba'i University, Tehran, Iran

² Department of English and Literature, Hakim Sabzevari University, Sabzevar, Iran

*Corresponding Author's Email: torkchinmatin.education@gmail.com

KEY TERMS

Language Teachers
Professional Identity
Work-related Elements
Organizational Elements
Personal elements

ABSTRACT

Numerous investigations have been carried out on the significance of language teachers' professional identity in recent years. As a crucial construct in educational psychology, professional identity has a critical role as teachers' practices and their educational approaches are formed by their identity. Several underlying constructs may affect language teachers' professional identity. This paper intends to review quantitative studies published in recent years. In the present review paper, the objective is to focus on the related studies that investigated the constructs that affect language teachers' professional identity. The underlying constructs were categorized into two sub-classes of personal elements and work-related and organizational elements. The former elements included autonomy, self-efficacy, critical thinking, and self-esteem, focusing on the involvement of personal factors. The latter class entailed burnout, job satisfaction, and other minor elements related to the work milieu and associated factors. The paper also presents a dearth of research and the gaps associated with professional identity investigation in language teaching and educational contexts.

ARTICLE TYPE

Review Paper

Received:	21 March 2025
Revised:	3 June 2025
Accepted:	9 June 2025
Published Online:	1 July 2025

© The Authors 2025

1. Introduction

Many investigations have been carried out on the language teachers' professional identity (LTPI) significance in language education since teachers have played fundamental roles in educational procedures (Awaje & Amaha, 2022; Dilek & Altas, 2022; Namaziandost et al., 2022; Nazari et al., 2021). According to some investigators, the research on the very scope lacks a clear description of the notion of LTPI (Karaolis & Philippou, 2019; Richardson & Watt, 2018). LTPI can be considered a trait formed

within particular social and pedagogical settings and can be promoted in a lengthy, stable, and successive procedure (Kalali Sani et al., 2022). Simply put, LTPI is formed by a series of emotional, personal, social, and contextual aspects that are responsible for the creation of self-perception and the roles they have in the classroom (Kang et al., 2022; Pishghadam et al., 2022). It can be defined as how teachers regard themselves as teachers, depending on having deeper insights into their ongoing relations with their educational environment (Beijaard et al., 2004). Furthermore, identity is conceptualized as how individuals view themselves and how others perceive them (Pennington & Richards, 2016).

Teachers who work on their effectiveness, influence, and self-directedness to attain more knowledge and skills are more inclined to develop their professional identity (PI), which is indispensable for their teaching career (Ivanova & Skara-Mincelne, 2016). As for student teachers, Izadinia (2013) believes that the social groups they are engaged in can affect their identity, which is viewed as a social construct and formed in social contexts. As college teachers are involved in the interaction and equilibrium of individual beliefs and norms process, their PI can be constructed and reconstructed, and they can be favorably impacted by their academic learning experience (Caihong, 2011). Recent studies have indicated that LTPI could contribute to their success positively, and the importance of PI should be highlighted by teacher education programs for developing instructors' performance (Derakhshan et al., 2020). Similarly, the ultimate purpose of a teacher education program is to elevate PI based on the Vygotskian outlook (Van Huizen et al., 2005).

Recent research has revealed a strong interrelationship between teachers' decision-making in educational contexts, especially within the classroom, and their identity, which can shape the essence of their profession (Bullough, 2011). Furthermore, EFL teachers' teaching experience influences PI impressively, and also when teacher trainers intend to set up some teacher training courses, they are required to consider the useful features of PI (Parsi & Ashraf, 2020). Ultimately, many contextual and individual factors impact LTPI, such as experience, beliefs, values, motivations, attributes, social interactions, socio-political context, culture, teacher education, school culture, and professional development (Derakhshan et al., 2020; Dilek & Altas, 2022; Motallebzadeh & Kazemi, 2018; Richardson & Watt, 2018).

In many current studies, a great number of significant factors have been identified that can impact LTPI very impressively, including teachers' autonomy (Dilek & Altas, 2022; Derakhshan et al., 2020), critical thinking (Moslemi & Habibi, 2019; Parsi & Ashraf, 2020; Sheybani & Miri, 2019), self-esteem (Motallebzadeh & Kazemi, 2018), burnout (Lu et al., 2019), and job satisfaction (Awaje & Amaha, 2022; Butakor et al., 2021). This review paper collects and reviews studies that scrutinize the relationship between LTPI and several related constructs, including job satisfaction, self-esteem, work engagement, self-efficacy, occupational commitment, critical thinking, autonomy, and burnout.

2. Theoretical Frameworks Related to Professional Identity

It is claimed that synthesizing theories related to LTPI is complex and confusing owing to its nature (Awaje & Amaha, 2022), and PI is regarded as a construct with different aspects (Richardson & Watt, 2018). Theories are the grounding aspects of a research study and several theories have been suggested to be associated with the PI as community of practice (CoP), sociocultural theory, activity theory, positioning theory, cultural production theory, practice theory, the dynamic system model of role identity (DSMRI), dialogical-self theory (Ahmad et al., 2019), and social cognitive theory (Marlow,

2009). Three main theories, namely CoP, socio-cultural theory, and activity theory, are primary in the scope of LTPI investigations that provide different visions toward PI.

Professional identity could be considered through social theories such as the Community of Practice. According to Awaje and Amaha (2022), this theory was developed by Wenger, and accordingly, this community refers to groups of individuals sharing common objectives, skills, and work milieu. They maintain that this theory is labeled as a social theory and the pair of learning and identity formation is not separable (Awaje & Amaha, 2022). A community of practice must be differentiated from communities that share the same interests or geographical boundaries due to the lack of a shared practice (Wenger, 1998). This theory is claimed to be the primary alternative to studying LTPI (Ahmad et al., 2019).

Learning and society are not separable as “we are social beings” and “this fact is a central aspect of learning” (Wenger, 2009, p.210). As teachers are one of the main learning agents, LTPI can be inspected through social theories. As Ahmad et al. (2019) mentioned, PI links with social constructivism and posited cognitive advancement would be enhanced with the zone of proximal development (ZPD) involvement with a person with more knowledge. They append that PI would be regarded as a cultural and social element that teachers would advance in their identity development by interaction with individuals under the dominance of ideas related to socio-cultural theory (Ahmad et al., 2019). As Richardson and Watt (2018) mentioned, teacher identity is rooted in social and cultural practices and it would be regarded as unlikely in distinct cultures of different communities. In addition, according to constructivist theories, identity is a dynamic trait formed in social interactions and changes in context (Pishghadam et al., 2022).

Vygotsky established Activity theory which is an extended theory of social constructivism (Ahmad et al., 2019, p.5). As Smit et al. (2010) mentioned, LTPI can be reasoned by activity theory as there are associations and interactions between all aspects of an activity system. It is stated that teachers are active representatives of self-growth and development, and there is mostly no tendency to act fully in their selected settings. It is also added that the class environment cannot be separated from the outside world as all parts of an activity system are related to each other because the environmental effect is evident (Smit et al., 2010). In the view of activity theory, it is worth mentioning that the subject refers to teachers who precede the activity with their standpoints, the object relates to the grounds for providing learners with practices and PI, “teaching approaches, methods, and methodologies” are associated with tools, all involved people in the course of teaching and learning are referred to as community, rules are labeled as yardsticks that systemize activity, and division of labor is related to the process of task distribution in a “hierarchical structure” among teachers (Ahmad et al., 2019, p. 5)

3. Elements of Professional Identity in Education

Professional identity cannot be addressed with rigid constructs, but researchers have proposed several traits to be related to the former (Ahmad et al., 2019). It could be analyzed through several constructs that have direct or indirect relationships with the former and have been investigated through various traits that would be different from study to study. The sub-elements of LTPI that are included in this review are autonomy, burnout, self-efficacy, critical thinking, and self-esteem as major constructs. However, there are minor elements that are tied with the PI or other constructs as experience, work engagement, occupational commitment, perception of the context, job stress, and motivation.

3.1 Personal Elements

In the scope of personal elements, the study outlines autonomy, critical thinking, and self-esteem. One or variant definitions of the construct are presented, followed by the studies carried out with a central focus on each construct. The participants, data collection samplings, data analysis methods, and the results are included in a brief presentation.

Autonomy has been characterized by Husband and Short (1994) as having the capability to determine the schedule and procedure of the work freely. Teacher autonomy can be defined as the expertise to form fine “skills, knowledge, and attitudes” (Smith, 2003, p. 9). This construct is addressed as the main factor in the work milieu that affects PI and job satisfaction (Strong & Yoshida, 2014). Autonomy is believed to be an acquired construct rather than intrinsic (Reich, 2002) and is claimed to be complex and contradictory (Pitt, 2010).

Motallebzadeh and Kazemi (2018) conducted a study on the connection between LTPI and self-esteem. This inquiry utilized convenience sampling with 224 EFL teachers and SPSS software for the data analysis. The SEM analysis revealed that sub-constructs of self-esteem were predictors of PI significantly and positively. All constitutions of self-esteem comprise satisfaction, knowledge, commitment, adaptation, and communication. A recent inquiry intended to scrutinize the connection between Iranian EFL LTPI, self-efficacy, and critical thinking skills in teaching procedures. Furthermore, 75 EFL teachers took part in the inquiry with an availability sampling type. For the analysis, SPSS was employed to calculate Pearson correlation and simple linear regression through ANOVA (analysis of variance models). The present inquiry's findings demonstrate a strong positive connection between the EFL LTPI, self-efficacy, and critical thinking adeptness. The ANOVA test revealed that the EFL LTPI is a predictor of self-efficacy and critical thinking adeptness (Moslemi & Habibi, 2019).

Critical thinking, as another factor, has been defined under the two lines of thought: solid defining of the concept and field-dependent defining. Sheybani and Miri (2019) attempted to promote an EFL LTPI relationship understanding with critical thinking. They concluded that LTPI correlates significantly with critical thinking, and also critical thinking positively impacts all three components of PI: subject matter field, didactical field, and pedagogical field (Sheybani & Miri, 2019).

A recent inquiry encompassed 190 EFL teachers. Pearson correlation was employed with SPSS, and structural equation modeling (SEM) was utilized with AMOS. The findings revealed that LTPI and autonomy efficiently predict their success. Additionally, they discovered significant positive relationships between LTPI, autonomy, and success (Derakhshan et al., 2020).

Parsi and Ashraf (2020) carried out an inquiry regarding the connection between EFL teachers' critical thinking and PI. This study subsumed 120 EFL teachers with convenience sampling. The study employed SPSS with Pearson correlation and multiple regression. The findings disclosed a strong positive connection between teachers' critical thinking and PI, which implies that as critical thinking is enhanced, PI is consequently improved. Furthermore, the results of multiple regression analysis indicated that the teaching experience of EFL teachers served as a more effective predictor of their PI.

A recent study was carried out on the connection between LTPI, self-esteem, and job satisfaction. Cluster and availability sampling were applied, and the total number of participants was 94 EFL teachers. In addition, Stata and analysis of moment structures (AMOS) were utilized in the data

analysis phase. The Pearson Correlation Coefficients indicated that self-esteem and job satisfaction are correlated to PI positively and significantly. Moreover, the results of SEM indicated that self-esteem and job satisfaction predicted teachers' PI significantly and positively. They maintained that self-esteem has an indirect mediating role between job satisfaction and LTPI (Awaje & Amaha, 2022).

Previous research aimed to examine how pre-service teachers' self-efficacy and learning engagement influence the connection between professional identities and experiences. A number of 309 pre-service teachers engaged in the study, and questionnaires were used to measure the variables. Structural equation modeling was utilized for the data analysis. A significant connection existed between learning engagement, PI, self-efficacy, and teaching internship. In a parallel fashion, learning engagement and self-efficacy both partially acted as a mediators in the connection between PI and teaching internship. The relationship between PI and teaching internship was also successively mediated by self-efficacy and learning engagement (Cai et al., 2022).

Dilek and Altas' (2022) investigation concerned the Turkish EFL LTPI relationship with their autonomy. The inquiry subsumed 250 with convenience sampling. Data analysis employed Pearson correlation with SPSS. The investigation results demonstrated a significant, positive, and moderate connection between PI and teachers' autonomy. Furthermore, they identified some elements, including sex, academic position, and job experience that could modify the association between LTPI and autonomy.

In recent research, Ding et al. (2022) intended to scrutinize the connection between EFL LTPI and critical thinking in a Chinese context. This study comprised 274 teachers, and their responses were elicited through convenience sampling. For data analysis, hierarchical linear regression was employed for direct hypotheses, and Hayes PROCESS models were used for mediation, moderation, and mediation moderation analysis. Hayes PROCESS model is a tool developed for analyzing complex relationships between variables, particularly in the context of mediation and moderation. The findings revealed that critical thinking mediates the relationship between LTPI and employee success. The findings of the moderated mediation analysis indicated the mediating function of critical thinking in the connection between LTPI and leaders' motivational language on teachers' success. Furthermore, related results denoted a positive connection between LTPI and critical thinking, leading to success impressively. Research measured the effect of critical thinking, grit, resilience, and self-efficacy perceptions on LTPI among Iranian EFL instructors. The study comprised 437 EFL teachers, and data analysis was carried out with confirmatory factor analysis (CFA) and SEM utilizing LISREL. In conclusion, they figured out that critical thinking and self-efficacy significantly affected LTPI (Namaziandost et al., 2022).

An inquiry is conducted on the connection between LTPI and professional development, with the mediating function of self-esteem, using 2668 special education teachers as participants. A random sampling was applied in the study, and IBM SPSS and the PROCESS macro program were utilized. The result of the connection between self-esteem and PI indicated a significant positive correlation. In addition, the findings revealed the mediating function of self-esteem on the connection between LTPI and professional development (Pi et al., 2024).

3.2 Work-related and Organizational Elements

In the scope of work-related and organizational elements, the paper describes job satisfaction and burnout. Alternative explication of the construct's notion is given, followed by the investigations

conducted with the primary focus on the constructs. The participants, data collection samplings, data analysis methods, and the findings are presented below.

Job satisfaction is another factor under this category that is a pleasant or aversive feeling of evaluating the job status and experience (McAllister et al., 2017; Troesch & Bauer, 2017). Furthermore, job satisfaction's effect modulates burnout among teachers, and job satisfaction correlates negatively with burnout (Chen et al., 2020; Lu et al., 2019). Furthermore, several investigations indicated that job satisfaction correlates positively with PI (Lu et al., 2019). Lu et al. (2019) conducted an investigation on the effect of job satisfaction on the relationships between LTPI and burnout. The responses were elicited from 267 student teachers through convenience sampling. Moreover, SPSS and Mplus as software, Pearson correlation, and SEM as data analysis methods were employed in the inquiry. The results of the Pearson correlation revealed a negative connection between LTPI and burnout, as well as job satisfaction and burnout. Inversely, job satisfaction correlated positively with LTPI. SEM results revealed that LTPI and job satisfaction were predictors of burnout, and job satisfaction partially acted as a mediator on the influence of LTPI on burnout.

As Chen et al. (2020) mentioned, job satisfaction mediates the connection between burnout and LTPI. The study comprised 483 Chinese university teachers. Additionally, SPSS and AMOS were utilized to carry out a common method bias test, descriptive statistics, and mediating effect analysis. The study reported that LTPI and job satisfaction significantly and negatively predict burnout. Furthermore, job satisfaction moderated the connection between LTPI and burnout. Burnout could be synthesized as having an opposing role concerning LTPI. Later, Butakor et al. (2021) inspected the relation between job satisfaction, work engagement, and PI. With a total sample size of 260 teachers, responses were elicited through multi-stage sampling. Exploratory factor analysis (EFA), SEM, and univariate statistical analyses were utilized to analyze the data. Moreover, SPSS was employed for EFA, and Mplus was utilized for SEM analysis. Related results indicated that job satisfaction acted as a mediator in the connection between the teacher's emotional intelligence and PI. They maintained that emotional intelligence would affect PI positively in a direct manner and through job satisfaction in an indirect manner.

Another recent research (Lin et al., 2022) was carried out to determine the connection between job satisfaction, work engagement, burnout, and PI. It was carried out with 3147 elementary and middle school teachers through random sampling, and SPSS was utilized for data analysis. Descriptive statistics and correlation analyses were applied to the main constructs. In addition, the SPSS PROCESS macro was employed for mediation and moderation analysis with 5000 bootstrapped samples. SPSS PROCESS macro is a tool developed for advanced statistical analyses focusing on mediation or moderation. The results indicated the mediating function of work engagement between PI and burnout. The connection between LTPI and burnout was negative, and the same result was found for the connection between burnout and work engagement, while the connection between LTPI and work engagement was a positive one. Furthermore, LTPI was recognized as the predictor of burnout.

Burnout is addressed as a result of sustained tension and stress upon teachers (Chen et al., 2020). It is claimed that teachers' burnout initiates with the student-teaching phase of their careers

(Hamman et al., 2007). There could be some reasons, such as teachers' characteristics in managing the classroom that would be responsible for teachers' burnout (Reichl, 2014) or excessive amounts of stress in the work milieu (Liu & Onwuegbuzie, 2014). Burnout and PI correlate negatively according to several investigations.

Burnout is investigated by Li et al. (2023) with the mediating function of work engagement and the moderating function of self-efficacy and perceived organizational support. This study was carried out on 3147 kindergarten, elementary, and middle school teachers. SPSS and ordinary least squares (OLS) were utilized to carry out the analysis. They reported that the connection between LTPI and burnout was partially mediated by work engagement. In addition, LTPI had the largest predictive influence on burnout via work engagement. Another research study was conducted on the connection between LTPI, burnout, work engagement, and psychological capital. A sum of 3147 primary and secondary school teachers attempted to answer the questionnaire, and SPSS was utilized. The findings indicated that teachers' PI, work engagement, and psychological capital correlated with burnout significantly in a negative way. Work engagement and psychological capital partially mediated the connection between LTPI and burnout (Zhang et al., 2024). The studies on LTPI, including constructs, designs, participants, sampling types, and relationships, are presented below in Table 1.

Table 1

A Review of Studies on Professional Identity

Author	Year	constructs	Design	Participants	Sampling	Relationship
Motallebzadeh & Kazemi	2018	Self-esteem	Quantitative survey	224 EFL Teachers	Convenience	Positive
Lu et al.	2019	Burnout Job satisfaction	Quantitative survey	267 Student Teachers	Convenience	LTPI> Negative Moderate burnout LTPI> Positive
Moslemi & Habibi	2019	Self-efficacy	Quantitative survey	75 EFL Teachers	Availability	Positive
Sheybani & Miri	2019	Critical Thinking	Quantitative survey	259 EFL teachers	Convenience	Positive
Chen et al.	2020	Critical thinking	Quantitative survey	483 Chinese university teachers	-	Positive
Derakhshan	2020	Autonomy	Quantitative survey	190 Iranian EFL teachers	Convenience	Positive
Parsi & Ashraf	2020	Critical Thinking	Quantitative survey	120 EFL teachers	Convenience	Positive
Butakor et al.	2021	Job satisfaction Job satisfaction	Quantitative survey	260 Teachers	Multi-Stage	Moderate Positive
Awaje & Amaha	2022	Job satisfaction Self-esteem	Cross-sectional	94 EFL Teachers	Cluster & Availability	Positive Positive

Cai et al.	2022	Self-efficacy	correlational research design	309 Preservice Teachers	Stratified random	Positive, mediate
Dilek & Altas	2022	Autonomy	Quantitative survey	250 Turkish EFL teachers	Convenience	Positive
Ding et al.	2022	Critical Thinking	Time-lag	274 employees	Convenience	Mediate
Han et al.	2022	Job satisfaction	Quantitative survey	648 Primary and Secondary Teachers	Convenience	LTPI> Positive
Lin et al.	2022	Burnout	Cross-sectional	3147 Primary and Secondary Teachers	Simple random	Negative
Namaziandost et al.	2022	Critical Thinking	Quantitative survey	437 EFL teachers	Random	Positive
Li et al.	2023	Burnout Self-efficacy Work engagement	Quantitative survey	3,147 school teachers	No access	LTPI> Negative Moderate Mediate
Pi et al.	2024	Self-esteem	Cross-sectional	2668 special education teach	Random	Positive, mediate
Zhang et al.	2024	Burnout	Descriptive correlational	3147 primary and secondary school teachers	No access	Negative

4. Discussion and Conclusion

The objective of this review was to outline the studies concerning LTPI and related constructs. In Table 1, the studies regarding LTPI are presented. Primarily, PI impacts teachers' performance and learners' learning processes crucially. Based on the conducted review, it was revealed that the designs of studies investigating LTPI were quantitative surveys, cross-sectional, time-lag, and correlational. The types of sampling in such studies were cluster, availability, multistage, convenience, and random sampling.

A strong PI among teachers can likely boost the educational system's effectiveness, and by this means, it is essential to invest on teachers' self (Awaje & Amaha, 2022) as they are agents who can contribute to the success of education (Cheng, 2021). Teachers are endowed with a crucial role in education by providing opportunities for their learners' nurturing (SoodmandAfshar & Hamzavi, 2017). Teacher identity is regarded as a multidimensional construct with an intrinsic association with teachers' learning and cognition (Pishghadam et al., 2022). Teacher identity is posited to be fluid, context-laden, and indiscriminate (Yuan & Zhang, 2020) and is labeled as a blurred and indistinct phenomenon (Namaziandost et al., 2022). Furthermore, LTPI can be characterized by interactions within communities of practice and the advancement of time (Namaziandost et al., 2022).

Conceptualizing LTPI more precisely, its relationship with some particular constructs must be scrutinized. LTPI is characterized by many personal constructs and work-related and organizational elements that interplay to build self-perception. Regarding personal factors, teachers

with a deep commitment to their career are likely to possess a strong personal identity, which is important for elevating their engagement and job satisfaction (Pi et al., 2024). Furthermore, the commitment becomes stronger by observing students' academic progress and improvement, which leads to intrinsic motivation for teachers (Wu et al., 2024). The present review paper focused on the related personal constructs including autonomy, self-efficacy, critical thinking, and self-esteem.

Language Teacher identity also relates to work-related and organizational constructs associated with contextual and social factors. Regarding contextual factors, teachers' perceptions of roles and responsibilities can be affected by changes in the educational system (Schiepe-Tiska et al., 2021). Educators can take responsibility for their professional identities in contexts that empower autonomy and teachers' practical knowledge (Pishghadam et al., 2022), while resistance and a distinct sense of identity that obstruct their professional growth may become emergent as a result of repeated and inadequately implemented changes (Jiang et al., 2021). Furthermore, social factors such as perceived support from colleagues and the community influence LTPI (Simon, 2024). In addition, supported teachers probably develop a more constant and positive PI, resulting in their commitment to learning (Lai et al., 2024). The underlying constructs of work-related and organizational types reviewed in this study are burnout, job satisfaction, occupational commitment, and work engagement.

According to Derakhshan (2020) and Dilek and Altas (2022), autonomy correlates positively with LTPI. It is reported that burnout correlates negatively with LTPI in a correlational (Chen et al., 2020; Lin et al., 2022; Lu et al., 2019; Zhang et al., 2024). Additionally, LTPI correlates negatively with burnout (Li et al., 2023). The studies on critical thinking and LTPI indicate a positive correlation between both constructs (Ding et al., 2022; Moslemi & Habibi, 2019; Namaziandost et al., 2022; Parsi & Ashraf, 2020; Sheybani & Miri, 2019). Related studies on job satisfaction and LTPI are abundant; they indicated a positive relationship between these traits (Awaje & Amaha, 2022; Butakor et al., 2021; Han et al., 2022; Lu et al., 2019). Furthermore, research indicated the moderating role of job satisfaction on LTPI and burnout (Chen et al., 2020). Self-efficacy correlated positively with LTPI in quantitative studies (Butakor et al., 2021; Moslemi & Habibi, 2019); it has a mediating function between LTPI and teaching internship (Cai et al., 2022), and a moderator function between LTPI and burnout (Li et al., 2023). Self-esteem was reported to positively correlate with LTPI (Awaje & Amaha, 2022; Motalebzadeh & Kazemi, 2018) and with a mediating function between LTPI and professional development (Pi et al., 2024). As Butakor et al. (2021) mentioned, work engagement has a positive correlation with LTPI, while in Li and colleagues' (2023) study, work engagement has a mediating function in the connection between LTPI and burnout.

Among various constructs investigated concerning LTPI, some are primarily focused in this review such as job satisfaction, work engagement, burnout, and critical thinking. Other constructs like autonomy, educational background, self-esteem, and self-efficacy could be investigated more to come to a richer conclusion. There are new approaches to PI in other fields of study, such as clinical, nursing, engineering, and studies related to general practitioners. In the aforementioned fields of study, constructs like career support, workforce policies, public image, self-understanding, turnover intention, age, working overtime, and occupational development opportunities are investigated, and one potential gap could be the lack of investigations on these traits in the very scope of LTPI. In addition, there are various theories applied or referred to in LTPI that could be the grounding frameworks for newer approaches to LTPI or newer aspects of the construct to be found. Therefore, future studies concerning LTPI could be focused on investigating other influential elements.

References

- Ahmad, H., Shah, S. R., Latada, F., & Wahab, M. N. (2019). Teacher identity development in professional learning: an overview of theoretical frameworks. *Bulletin of advanced English studies*, 3(1), 1–11. <https://doi.org/10.31559/baes2019.3.1.1>
- Awaje, A. H., & Amaha, H. B. (2022). The relationship between EFL teachers' professional identity, professional self-esteem, and job satisfaction. *Bahir Dar Journal of Education*, 22(2), 89–105.
- Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education*, 20(2), 107–128. <https://doi.org/10.1016/j.tate.2003.07.001>
- Bullough, R. V. (2011). Ethical and moral matters in teaching and teacher education. *Teaching and Teacher Education*, 27(1), 21–28. <https://doi.org/10.1016/j.tate.2010.09.007>
- Butakor, P. K., Guo, Q., & Adebajji, A. O. (2021). Using structural equation modeling to examine the relationship between Ghanaian teachers' emotional intelligence, job satisfaction, professional identity, and work engagement. *Psychology in the Schools*, 58(3), 534–552. <https://doi.org/10.1002/pits.22462>
- Cai, Z., Zhu, J., & Tian, S. (2022). Preservice teachers' teaching internship affects professional identity: Self-efficacy and learning engagement as mediators. *Frontiers in Psychology*, 13, 1070763. <https://doi.org/10.3389/fpsyg.2022.1070763>
- Caihong, H. (2011). Changes and characteristics of EFL teachers' professional identity: the cases of nine university teachers. *Chinese Journal of Applied Linguistics*, 34(1). <https://doi.org/10.1515/cjal.2011.001>
- Chen, H., Liu, F., Pang, L., Liu, F., Fang, T., Wen, Y., Chen, S., Xie, Z., Zhang, X., Zhao, Y., & Gu, X. (2020). Are you tired of working amid the pandemic? The role of professional identity and job satisfaction against job burnout. *International Journal of Environmental Research and Public Health*, 17(24), 9188. <https://doi.org/10.3390/ijerph17249188>
- Cheng, L. (2021). The Implications of EFLESL teachers' emotions in their professional identity development. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.755592>
- Derakhshan, A., Coombe, C., Arabmofrad, A., & Taghizadeh, M. (2020). Investigating the effects of English language teachers' professional identity and autonomy in their success. *Issues in Language Teaching*, 9(1), 1–28. <https://doi.org/10.22054/ilt.2020.52263.496>

- Dilek, E., & Altas, B. (2022). The relationship between EFL teachers' professional identity and professional autonomy. *Asian Journal of Education and Training*, 8(1), 22–33. [https://doi.org/10.20448/edu.v8i1.3806"1.3806](https://doi.org/10.20448/edu.v8i1.3806)
- Ding, F., Liu, X., Amin Abdalla, A., Latif Khan, M., & Akram, F. (2022). The link between English foreign language teacher's professional identity and their critical thinking that leads to teacher's success in the Chinese context: Leaders motivational language as a moderator. *Frontiers in Psychology*, 13, 983215. <https://doi.org/10.3389/fpsyg.2022.983215>
- Hamman, D. (2007). Efficacy and pedagogical interaction in cooperating and student teacher dyads. *Journal of Classroom Interaction* Vol, 42(55), 41–42. <https://www.jstor.org/stable/23869448>
- Husband, R. E., & Short, P. M. (1994). Interdisciplinary teams lead to greater teacher empowerment. *Middle School Journal*, 26(2), 58–60. [https://doi.org/10.1080/00940771.1994.11494412"00940771.1994.11494412](https://doi.org/10.1080/00940771.1994.11494412)
- Ivanova, I., & Skara-MincLne, R. (2016). Development of professional identity during teacher's practice. *Procedia - Social and Behavioral Sciences*, 232, 529–536. [https://doi.org/10.1016/j.sbspro.2016.10.073"2016.10.073](https://doi.org/10.1016/j.sbspro.2016.10.073)
- Izadinia, M. (2013). A review of research on student teachers' professional identity. *British Educational Research Journal*, 39(4), 694–713. [https://doi.org/10.1080/01411926.2012.679614"01411926.2012.679614](https://doi.org/10.1080/01411926.2012.679614)
- Jiang, H., Wang, K., Wang, X., Lei, X., & Huang, Z. (2021). Understanding a STEM teacher's emotions and professional identities: A three-year longitudinal case study. *International Journal of STEM Education*, 8, 1–22. <https://doi.org/10.1186/s40594-021-00309-9>
- Kalali Sani, S. F., Motallebzadeh, K., Khodabakhshzadeh, H., & Zeraatpisheh, M. (2022). On the Iranian English as foreign language novice and experienced teachers' attributional styles and professional identity. *Frontiers in Psychology*, 12, 823815. [https://doi.org/10.3389/fpsyg.2021.823815"2021.823815](https://doi.org/10.3389/fpsyg.2021.823815)
- Kang, M., Shen, Q., & Zheng, Y. (2022). LOTE (Languages Other than English) Teachers' emotions and professional identity in response to educational reforms: A social-psychological perspective. *Sustainability*, 14(17), 10788. <https://doi.org/10.3390/su141710788>
- Karaolis, A., & Philippou, G. N. (2019). Teachers' professional identity. *Affect and mathematics education: Fresh perspectives on motivation, engagement, and identity*, 397–417.
- Lai, L., Shi, W., Xie, Y., & Zhao, J. (2024). Factors influencing the generation of teachers' emotions. *Frontiers in psychology*, 15, 1392965. <https://doi.org/10.3389/fpsyg.2024.1392965>

- Li, X., Xiao, W., Sun, C., Li, W., & Sun, B. (2023). Does burnout decrease with teacher professional identity among teachers in China? *Journal of Career Development*, 50(5), 983–996. <https://doi.org/10.1177/08948453221138937>
- Lin, Y., Ameyaw, M. A., Zhang, Q., Sun, B., & Li, W. (2022). The relationship between teacher professional identity and burnout amid the pandemic: A moderated mediation model. *Frontiers in Public Health*, 10, 956243. <https://doi.org/10.3389/fpubh.2022.956243>
- Liu, S., & Onwuegbuzie, A. J. (2014). Teachers' motivation for entering the teaching profession and their job satisfaction: A cross-cultural comparison of China and other countries. *Learning Environments Research*, 17(1), 75–94. <https://doi.org/10.1007/s10984-013-9155-5>
- Marlow, M. (2009). Supporting teacher professional identity through mentoring activities. *Research in Higher Education Journal*, 2, 1. Maslach, C., & Leiter, M. P. (2023). Burnout. In *Encyclopedia of Mental Health* (pp. 315–322). Elsevier. <https://doi.org/10.1016/B978-0-323-91497-0.00128-4>
- McAllister, C. P., Harris, J. N., Hochwarter, W. A., Perrewé, P. L., & Ferris, G. R. (2017). Got resources? A multi-sample constructive replication of perceived resource availability's role in work passion–job outcomes relationships. *Journal of Business and Psychology*, 32(2), 147–164. <https://doi.org/10.1007/s10869-016-9441-1>
- Moslemi, N., & Habibi, P. (2019). The relationship among Iranian EFL teachers' professional identity, self-efficacy and critical thinking skills. *HOW*, 26(1), 107–128. <https://doi.org/10.19183/how.26.1.483>
- Motallebzadeh, K., & Kazemi, B. (2018). The relationship between EFL teachers' professional identity and their self-esteem. *Cogent Education*, 5(1), 1443374. <https://doi.org/10.1080/2331186X.2018.1443374>
- Namaziandost, E., Heydarnejad, T., & Saeedian, A. (2022). Language teacher professional identity: The mediator role of L2 grit, critical thinking, resilience, and self-efficacy beliefs. *Iranian Journal of Applied Language Studies*, 14(2), 107–130. <https://doi.org/10.22111/ijals.2022.7486>
- Nazari, M., Zohoorian, Z., & Faravani, A. (2021). Patterns of teaching empowerment and classroom management among successful EFL teachers. *International Journal of Research in Teacher Education*, 12(2), 36–52.
- Parsi, G., & Ashraf, H. (2020). The relationship among EFL teachers' critical thinking, professional identity and teaching experience. *Journal of Applied Linguistics and Language Research*, 7(1), 144–154. Retrieved from www.jallr.com

- Pennington, M. C., & Richards, J. C. (2016). Teacher identity in language teaching: integrating personal, contextual, and professional factors. *RELC Journal*, 47(1), 5–23. <https://doi.org/10.1177/0033688216631219>
- Pi, Y., Ma, M., Hu, A., & Wang, T. (2024). The relationship between professional identity and professional development among special education teachers: a moderated mediation model. *BMC psychology*, 12(1), 1–14. <https://doi.org/10.1186/s40359-024-02075-z>
- Pishghadam, R., Golzar, J., & Miri, M. A. (2022). A new conceptual framework for teacher identity development. *Frontiers in Psychology*, 13, 876395. <https://doi.org/10.3389/fpsyg.2022.876395>
- Pitt, A. (2010). On having one's chance: autonomy as education's limit. *Educational Theory*, 60(1), 1–18. <https://doi.org/10.1111/j.1741-5446.2009.00342.x>
- Reich, R. (2002). Opting out of education: Yoder, Mozart, and the autonomy of children. *Educational Theory*, 52(4), 445.
- Reichl, C., Wach, F.-S., Spinath, F. M., Brünken, R., & Karbach, J. (2014). Burnout risk among first-year teacher students: The roles of personality and motivation. *Journal of Vocational Behavior*, 85(1), 85–92. <https://doi.org/10.1016/j.jvb.2014.05.002>
- Richardson, P. W., & Watt, H. M. G. (2018). Teacher professional identity and career motivation: a lifespan perspective. In P. A. Schutz, J. Hong, & D. Cross Francis (Eds.), *Research on Teacher Identity* (pp. 37–48). Springer International Publishing. https://doi.org/10.1007/978-3-319-93836-3_4
- Schiepe-Tiska, A., Dzchaparkulova, A., & Ziernwald, L. (2021). A mixed-methods approach to investigating social and emotional learning at schools: Teachers' familiarity, beliefs, training, and perceived school culture. *Frontiers in psychology*, 12, 518634. <https://doi.org/10.3389/fpsyg.2021.518634>
- Sheybani, M., & Miri, F. (2019). The relationship between EFL teachers' professional identity and their critical thinking: A structural equation modeling approach. *Cogent Psychology*, 6(1), 1592796. <https://doi.org/10.1080/23311908.2019.1592796>
- Simon, E. (2024). Cultivating professional identity: The vital role of practical teaching experience for future educators. *Education Sciences*, 14(5), 439. <https://doi.org/10.3390/educsci14050439>
- Smit, B., Fritz, E., & Mabalane, V. (2010). A conversation of teachers: In search of professional identity. *The Australian Educational Researcher*, 37(2), 93–106. <https://doi.org/10.1007/BF03216924>

- Smith, M. (2003). Towards a professional identity and knowledge base: Is residential child care still social work?. *Journal of Social Work*, 3(2), 235-252. <https://doi.org/10.1177/14680173030032007>
- SoodmandAfshar, H., & Hamzavi, R. (2017). An investigation into the characteristics of Iranian EFL teachers of senior secondary schools and language institutes. *Iranian Journal of Language Teaching Research*, 5(1), 21-36. <https://doi.org/10.30466/ijltr.2017.20340>
- Strong, L. E. G., & Yoshida, R. K. (2014). Teachers' autonomy in today's educational climate: current perceptions from an acceptable instrument. *Educational Studies*, 50(2), 123-145. <https://doi.org/10.1080/00131946.2014.880922>
- Troesch, L. M., & Bauer, C. E. (2017). Second career teachers: Job satisfaction, job stress, and the role of self-efficacy. *Teaching and Teacher Education: An International Journal of Research and Studies*, 67(1), 389-398. <http://dx.doi.org/10.1016/j.tate.2017.07.006>
- Van Huizen, P., Van Oers, B., & Wubbels, T. (2005). A Vygotskian perspective on teacher education. *Journal of Curriculum Studies*, 37(3), 267-290. <https://doi.org/10.1080/0022027042000328468>
- Wenger, E. (1998). Communities of practice: Learning as a social system. *Systems thinker*, 9(5), 2-3.
- Wenger, E. (2009). A social theory of learning. In *Contemporary theories of learning* (pp. 217-240). Routledge.
- Wu, J., Ghayas, S., Aziz, A., Adil, A., & Niazi, S. (2024). Relationship between teachers' professional identity and career satisfaction among college teachers: role of career calling. *Frontiers in Psychology*, 15, 1348217. <https://doi.org/10.3389/fpsyg.2024.1348217>
- Zhang, Q., Li, W., Gao, J., Sun, B., & Lin, S. (2024). Teachers' professional identity and job burnout: The mediating roles of work engagement and psychological capital. *Psychology in the Schools*, 61(1), 123-136. <https://doi.org/10.1002/pits.23039>



Contents lists available at [JSLP](#)

Journal of Second Language Pedagogy

Journal homepage: <https://www.sanad.iau.ir/journal/jslp>

Secondary- level EFL Teachers' Perceptions toward Doing Research and Their Challenges in Research Engagement

Sara Noroozi¹, Effat Fathi², Gholamreza Zareian^{3*}, Mohammad Reza Amirian⁴

¹ Department of English, Ne.C. Islamic Azad University, Neyshabour, Iran

² Department of Education, Sabzevar

³ Hakim Sabzevari University

⁴ Hakim Sabzevari University

*Corresponding Author's Email: gzareian@hsu.ac.ir

KEY TERMS

Challenges in research engagement
Iranian EFL teachers
Reasons for research engagement
Research engagement
Teacher research

ARTICLE TYPE

Original Research Paper

Received:	1 March 2025
Revised:	22 June 2025
Accepted:	25 June 2025
Published Online:	1 July 2025

© The Authors 2025

ABSTRACT

Research engagement seems to be an important means for teachers to develop their professional competence and status in different aspects; thereby, the present study employs a mixed-methods design to examine English language teachers' engagement with and in research in the context of the secondary level EFL teachers in Iran. Using a Likert scale questionnaire followed by in-depth interviews, the researcher gathered empirical information from 100 English language teachers who teach at various lower-secondary and upper-secondary public high schools mostly located in the city of Sabzevar in Khorasan Razavi province. The objectives of the study were to investigate the extent of Iranian EFL teachers' engagement with/in research, the relationship between some related factors including EFL teachers' years of teaching experience, their qualifications, and genders on the extent of research engagement (RE), and also to explore the teachers' perceptions about the important reasons for being research engaged and the dominant challenges and barriers EFL teachers deal with in the context of the study. The findings indicated that the EFL teachers' engagement in/with research in the research context was not a frequent activity and most of them were engaged with/in research in a less than a moderate level; therefore, RE was a minority activity. EFL teachers also highlighted mostly external reasons for research engagement. The impeding factors were firstly related to lack of financial supports and economic matters and secondly to the barriers related to non-collaborative school culture in the context of the study. The study has some implications and also some suggestions for further studies.

1. Introduction

Teachers are one of the most important contributors to educational system, and their responsibility in schools extends beyond carrying out and delivering the curriculum. They also need to know how to recognize and solve problems that may occur within the classroom when they deliver the curriculum (Davies, 1995). It is a broad and important aspect of development not only for the teacher-researchers themselves but also for the colleagues with whom they share results in writing or at conferences or meetings and they can do research at all levels of education. Teacher research may mean different things to different educators. To some it may be engagement with research that pertains to reading and using research while for others it may be doing research and publishing the results (Borg, 2010; Kyaw, 2021). Typically, research is conducted because teachers want to have some questions answered--hence the development of the teachers' goals for the research. These questions "often develop gradually as teachers try to figure out why certain things are happening in their classrooms" (Hubbard & Power, 1993, p. 20).

Moreover, a qualified teacher should be aware of, and be able to respond to the direction of new development in teaching (Lewis & Munn, 1997). In other words, in modern world, effective teaching needs teachers engage in educational research in order to improve the standards of their teaching. At the present time, and especially in the developing countries, teachers are expected to follow educational research findings in order to enhance the quality of their teaching, and to solve problems that may come up in their classes (Everton et al., 2002). Moreover, research engagement has been increasingly seen as an important means for English language teachers to improve their professional competence in many contexts (Ahmed & Pervin, 2013; Atay, 2008; Borg, 2009). Consequently, English teachers in these contexts have been called on to play "a central and critical role in generating knowledge of practice by making their classrooms and schools sites for enquiry" (Cochran-Smith & Lytle 1999, p. 273).

Although teacher research engagement is not an old phenomenon in English language teaching (ELT), there are only a limited number of empirical studies related to high school teachers' understanding of research in comparison to university contexts (Bai & Millwater, 2011; Sippel & Sato, 2022) and relative to levels of research engagement (Mehrani & Behzadnia, 2013; Mehrani, 2015; Mehranirad & Behzadpoor, 2022), and research which suggested teachers how to promote teachers' research engagement (Allwright & Baily, 1991; Li & Xu, 2024; Nunan 1992; Freeman, 1998; Mehrani, 2015; Mehranirad, 2023). Furthermore, research engagement is often a challenging task for English teachers, especially for teachers, who are relatively underpaid and whose professional contributions receive little recognition in comparison with those of their counterparts in other educational settings (Allwright 1997). Mehranirad and Behzadpoor (2022) found that the barriers on their way of engagement in and with research are relevance of research findings to their needs, restrictive educational policies, lack of teacher-research collaboration, and insufficient technological facilities and resources. With this in mind, this study aims to uncover Iranian EFL teachers' beliefs about the reasons for research engagement, measure the extent of their engagement in research and identify barriers that prevent them to be engaged in research.

2. Methodology

2.1 Participants and Setting

The sample of the study comprises of 100 EFL teachers with an age range between 25-50 years, including male and female EFL teachers (60 female, 40 male) teaching in public junior or senior high schools and/or private language institutions. They held BA (70%) or MA (30%) degrees in TEFL, translation, and English literature. Teachers participated in the present work were selected through non-probability convenience or opportunity sampling procedure because of some practical aspects such as geographical proximity, availability and their voluntary willingness to participate in the study. Most of the teachers were working in junior high schools, senior high schools and/pre-universities and some of them from private language institutions in Sabzevar and some other cities in Khorasan Razavi province such as Shirvan and Neyshaboor and also from Tehran.

2.2 Instrumentation

To obtain the needed data for the present study, the researcher utilized the following instruments.

English Language Teachers' Conceptions of Research Questionnaire

As a means of data collection, a questionnaire can let large amount of data be collected efficiently, economically, and in a standardized manner (Dornyei & Taguchi, 2010). In addition, considering the accessibility to some geographically dispersed participants, a questionnaire can be easily administered through the internet access. Hence survey studies aim at describing the characteristics of a population the use of questionnaires is the main data collection procedure in surveys (Dornyei, 2007).

The questionnaire used here was adapted from Borg's (2009) study of English language teachers' conceptions of research. To make the questionnaire relevant to the context of the current study, the researchers modified the original one based on feedback from a panel of experts in applied linguistics. Borg's questionnaire includes six sections focusing on different themes of his study. Regarding some sections and also some questions and sub-questions of the original version unnecessary to the objectives of the present study, the researchers revised the questionnaire and deleted unrelated sections or items according to the experts' recommendations. For pilot-testing, it was distributed among a group of 20 EFL teachers who were similar to the study sample to measure the internal consistency of the questionnaire that yielded an acceptable Cronbach's alpha of 0.7.

The other important modification was the constructing of two Likert scale questionnaires based on the reasons of being research engaged and barriers of RE which had been identified in Borg's (2009) study. In addition, each of these two Likert scale questionnaires were divided into some subtitles according to the benefits and barriers of RE which were indicated in the literature (e.g., Atay, 2008; Allison & Carey 2007; Burton & Mican 1993; McKernan 1993; Borg 2003, 2009; Worrall, 2004; McKay 2009) and were classified by Borg (2009).

In this way, the final version of the questionnaire is composed of four sections. In section one there are five items to provide some demographic information necessary to the study, the second section of this instrument includes four close ended questions to measure the extent engagement in research, the third section of the questionnaire is a Likert scale questionnaire consisting of 16 items under two subtitles to infer the most salient and frequent reasons for doing or reading research, and finally the fourth section four is also a Likert scale consisting 24 items under five subcategories to classify and infer the dominant barriers to RE in the context of the study.

The subtitles of the third section included 16 items on reasons for the effectiveness of RE on teaching which are determined with 7 items (3, 4, 5, 8, 9, 10, 12) and reasons of RE for professional development in 9 items (1, 2, 6, 7, 11, 13, 14, 15, 16). The fourth section of the questionnaire composed of five sub sections to determine the obstacles to RE related to non-collaborative school culture which is measured by 4 items (40, 33, 34, 11), limitations in teachers awareness, beliefs, skills and knowledge to be inferred by 4 items (19, 2, 16, 12), limited resources with 5 items (13, 22, 23, 14, 17), demotivators with 7 items (32, 37, 39, 24, 29, 25, 9) and finally economic matters with 4 items (38, 36, 27, 28).

Semi-Structured Interviews

In spite of the popularity of the questionnaire in applied linguistics to collect qualitative data, this is not to overlook the drawbacks of the questionnaires in both their conventional and electronic formats, for example they are vulnerable to design flaws (Borg, 2013) and according to Bryman (2007), in comparison to interviews, questionnaires do not provide in-depth exploration of particular issues and they are more subject to respondents' fatigue. Therefore, to have a more in-depth exploration of issues and gather some reasons and barriers probably not being covered in the Likert scale questionnaire, the researcher conducted an interview stage through which 10 people of the respondents who completed the questionnaire were selected purposefully to participate in the interview.

To ask the respondents for interview, while completing the questionnaire teachers were asked if they would like to participate in a follow up interview to discuss the related issues, they would provide their name and contact details. After receiving respondents' completed questionnaires, a series of follow up questions were conducted to explore more detail or to exemplify some of the general issues. The questions of interview were structured according to the topics constructed in the questionnaire but the scope was more flexible to elaborate any matter more effectively. The interviews lasted an average of 30 minutes and according to the previous appointment they were administered face. The main interview questions were as follows:

1. Are you engaged in research?
2. If not, what prevents you from doing research?
3. What's the main reason for this problem? Personal, institutional, or any other?

2.3 Design of the Study

In the present study the researcher employs a mixed method deign. A mixed method study involves the mixing of quantitative and qualitative research methods or paradigm characteristics (Johnson and

Christensen, 2004), Therefore the design of the study adopted what Creswell(2003) calls a 'sequential explanatory strategy, which is a mixed methods design characterized by the collection and analysis of quantitative data followed by the collection and analysis of qualitative data, the priority typically given to the quantitative data and the two methods are integrated during the interpretation phase of the study '(p. 215). The quantitative data needed for the present study was generated via a Likert scale questionnaire and then the follow up interviews were conducted to collect the supporting qualitative data to achieve a fuller understanding of the target phenomenon.

2.4 Research Procedure

As mentioned above this study adopted a mixed methods strategy in data collection to enhance the quality of research. In this study by the use of both questionnaire and interview the researcher aims to 'examine the evidence from different sources to build a coherent justification for themes'(Creswell, 2003, p.196). To follow this method, the participants were asked either directly by the researcher or indirectly through the heads of English teachers' communities or other colleagues to fill up the questionnaire. Some relevant information about the purpose of the study was provided for the teachers before the distribution of the scales. In addition, confidentially assurance and in-advance acknowledgement for their cooperation in the study were given in an introduction at the first page of the questionnaire.

Some of the questionnaires were directly distributed to the participants. In some cases, the researcher attended schools or institutions to distribute the questionnaires. In one case the researcher attended the national conference of challenges on EFL teaching and learning holding in Hakim Sabzevari University and distributed some of the questionnaire to EFL teachers who participated in the conference. For those of the respondents who were geographically dispersed, soft copies of the questionnaires were sent to them via email.

After the submission of questionnaires, the researcher contacted with the respondents who accepted to interview, and make appointment to do face to face or via telephone interview. The researcher preferred face to face interview. The interview set up in the study is a semi- structured interview (Dornyei, 2007) in which "although there is a set of pre-prepared guiding questions and prompts the format is open-ended and the interviewee is encouraged to elaborate on the issues raised in an explanatory manner"(p. 136). Qualitative data were collected through interviews set up for 10 EFL teachers who announced their tendency to participate in the interview. Before starting the interview, the researcher briefly discussed the purpose of the interview and the use of audio recorder. In addition to record the interview via a voice recorder, the researcher takes note during the interviews.

3. Data Analysis

The present study intended to investigate the extent of IEFL teachers' research engagement and the relationship between EFL teachers' RE and their years of experience, their academic degrees and their gender. Moreover, an attempt was made to determine challenges teachers had in their research engagement. Therefore, in this section, the related descriptive and inferential results as well as the interview findings are presented and discussed.

The Extent of Research Engagement

The first objective of the study was to determine the frequency of doing research by the teachers. It is clear from Table 1 that 47% indicated that they do research sometimes, 36% rarely, 10% often, 2% always and 5% never. Therefore, it can be claimed that doing research is not a frequent activity among IEFL teachers and about 83% are engaged in doing research sometimes or rarely.

Table 1

Reported Frequency of Doing Research by EFL Teachers

Frequency	Number of Respondents	Percent of the Respondents
Always	2	2%
Often	10	10%
Sometimes	47	47%
Rarely	36	36%
Never	5	5%

Regarding the extent to which the respondents utilized the published studies and resources associated with TEFL, Table 2 shows that 56% of respondents said they sometimes read the published researches, 22% rarely, 15% often, 5% always, and 2% never studied. Therefore, it can be said that while a majority of respondents have a limited level of engagement with published research.

Table 2

Reported Frequency of Reading and Using Published Research

Frequency	Number of respondents	Percent of the respondents
Always	5	5%
Often	15	5%
Sometimes	56	56%
Rarely	22	22%
Never	2	2%

To determine whether there was a relationship between the years of the teaching experience and the extent of RE, it was found that there wasn't any significant difference between the respondents in this regard as displayed in table 3. In this part, the respondents were divided into three groups according to the years of teaching experience: the first group 0-9 years, the second one 10-19 years, and the last group at least 20 years.

Table 3*The Relationship between Years of Experience and the Extent of RE*

Test Statistics ^{a,b}	
	n6
Chi-Square	3.687
Df	2
Asymp. Sig.	.158
a. Kruskal Wallis Test	
b. Grouping Variable: group	

In order to investigate the effect of teachers' academic degrees on their research engagement, Mann-Whitney U test was run and the results (as shown in Table 4) revealed that there was a significant difference between the respondents' research engagement regarding their academic degrees.

Table 4*The Relationship between Teacher's Qualification and the Extent of RE*

Test Statistics ^a	
	n6
Mann-Whitney U	502.500
Wilcoxon W	1.282E3
Z	-5.917
Asymp. Sig. (2-tailed)	.001
a. Grouping Variable: n2	

To explore the relationship between gender and the extent of RE, another Mann-Whitney U test was run. The findings revealed that female EFL teachers had higher extent of RE in comparison to the male EFL teachers (Table 5).

Table 5*The Relationship between Gender and RE*

Test Statistics ^a	
	n6
Mann-Whitney U	830.000
Wilcoxon W	1.650E3
Z	-3.173
Asymp. Sig. (2-tailed)	.002
a. Grouping Variable: gender	

The Reasons for Doing Research

In studies conducted in different contexts (Borg & Liu 2013; Everton et.al 2002; Gao & Chow 2011; McNamara 2002; Worrall 2004) different reasons for doing research were explored and listed. In addition, some important benefits for teacher research (TR) were elaborated in the literature. The present study aimed to explore the most frequent reasons for being research engaged by IEFL teachers. The reasons were grouped into two categories: The reasons which dominantly express the influence of RE on teaching procedures and the reasons for EFL teachers' professional status.

A fundamental argument is that when teachers engaged with and in research and make pedagogical decisions informed by sound research evidence this will have a beneficial effect of both teaching and learning, additionally the evidence-based practice has been extensively described and debated in the literature (e.g., Davies, 1999; Elliot, 2002). According to Table 4.8 one of the significant reasons for teachers to be research engaged can be the beneficial impact that it has on the quality of teaching, in this regard the first part of the third section in the questionnaire aimed to measure the level of participants coincidence on some influencing factors of RE on teaching. As it is showed in Table 4.8 one of the fundamental reasons as indicated by the respondents was that RE "allows teachers to become more reflective, critical and analytical about their teaching behaviors in the classroom" (with 94% agreement). Two other salient reasons for RE with the same level of agreement (89%) were "doing research to find better ways of teaching" and "to contribute to renew and update information about language teaching". The reason of "improving teaching procedure" gained 82% agreement. The other reasons in this part received on average of 73% accordance on the part of respondents.

Table 6

The Reasons for Doing Research to Influence Teaching Quality

Statement	Strongly Disagree (%)	Disagree (%)	Don't know (%)	Agree (%)
allows teachers to became more reflective, critical.....	0	3	3	94
creates a problem-solving mindset that help teachers when....	1	3	15	81
develops teachers' capacity for autonomous professional judgments	2	2	25	71
Helps to find better ways of teaching	1	5	5	89
allows me to collaborate with colleagues	1	10	22	67
causes to improve my teaching procedure	1	3	14	82
contributes to update my information about language teaching	1	5	5	89

Research engagement has been increasingly seen as an important means for English language teachers to improve their professional competence in many contexts (Atay 2008; Borg 2009). Besides, teacher research has been advanced as an important means for in-service teachers to improve their professional competence (Gu and Wang 2006; You 2007), thereby a number of EFL teachers are engaged in research to develop their professional status. The concern of the second part in the third section of the questionnaire is to examine and measure the majority and minority of such reasons from the EFL teachers' points of view. Nine reasons for professional development were presented to the teachers in this part in the form of Likert scale. Among these nine statements (reasons) a considerable number of respondents (about 81%) agreed that "doing research reduces the feeling of frustration of their jobs", 80% of them agreed that RE "allows teachers to move out of a submissive position and be a curriculum inventor". The other salient reasons for being research engaged were "helping to solve professional problems" 80% and "improving teachers' instructional decision-making process" 75%. Other reasons in this part received at least the average of 57% agreement among respondents. Table 7 shows the results in more details.

Table 7*The Reasons for Doing Research for Professional Development*

Statement	Strongly Disagree (%)	Disagree (%)	Don't know (%)	Agree (%)
reduces teachers feeling of frustration and isolation	5	6	9	80
allows teachers to move out of a submissive position and be curriculum innovators	1	7	12	80
improves teachers instructional decision-making process	0	6	19	75
increases the professional status of teachers	4	15	26	55
helps empower teachers to influence their own profession	5	11	28	56
contributes to improvement of my department	7	19	27	47
it is good for my professional development	7	10	18	65
I feel that doing research is an important part of my job	2	12	24	62
helps me to solve my professional problems	0	5	15	80

The Barriers to Research Engagement

Quantitative Results

In comparison to the volume of empirical work which has been conducted in to teachers' research engagement in education, generally little research has been conducted in the field of English language teaching (Borg, 2003), Crooks and Arakaki (1999), McDonough & McDonough (1990) refer to some factors and barriers which hinder teachers research engagement. Additionally, the reasons for not doing research have been recognized in the related literature and also categorized in different subtitles (Borg & Liu, 2013).

Although some obstacles are of the same sorts in different context, one purpose of the present study was to recognize which of such barriers are the most common among the respondents of the Iranian context. Thus, the first part of the section four in the questionnaire consisted of four barriers under the title of "non- collaborative school culture". In general, 77% of the participants chose "agree" and "strongly agree" for the first item that said "the employers do not support them in doing research". The second noteworthy barrier for the respondents was "the lack of cooperation of the principals or department head for teachers to be research engaged" which 35% were agreed. The other two non-collaborative school culture obstacles seemed not to be very dominant for the respondents as displayed in Table 8.

Table 8

Barriers Related to Non-Collaborative School Culture

Statement	Strongly Disagree (%)	Disagree (%)	Don't know (%)	Agree (%)
My employer does not support me	1	3	19	77
Other teachers would not cooperate if I asked for their help	1	36	43	20
Learners would not cooperate if I did research in class	10	55	18	17
Principal or department head dose not cooperate me to do research	12	20	33	35

Considering the barriers related to the "limitations in teachers' awareness, beliefs, skills and knowledge" the most notorious obstacle was "the teacher's perception of research as an academic, large scale, statistical and technically difficult activity on which 63% of the respondents reported agreement. Another important obstacle with less than 50% agreement were: having not access to needed books and journals (43%), thinking that research publications are irrelevant to their teaching contexts (41%), having not enough knowledge about research method (40%), and finally a less dominant preventive factor in this part was having difficulty in understanding research reports with 19% agreement as shown in Table 9.

Table 9*Limitations in Teachers' Awareness, Beliefs, Skills and Knowledge*

Statement	Strongly Disagree (%)	Disagree (%)	Don't know (%)	Agree (%)
I do not know enough about research methods	14	35	11	40
I face difficulty in understanding research reports	18	41	22	19
I think of research as an academic, large scale, statistical and technically difficult activity	2	22	13	64
Most research publication are not relevant to my teaching context	6	38	15	41

"Limited resources" was the other important category of barriers with six influencing factors under it as shown in Table 10. Referring to the obtained results of "having not enough knowledge about how to access to digital libraries and web-based resources of research", 83% agreement on this item revealed its importance as one of the prominent barriers. Lack of ability to think of note-worthy topics for research received 50% agreement. The other important but not very dominant ones were "having no access to English resources such as books and journals" (37%), and "not allocating a particular time to do research as a part of teachers' yearly workload" (36%).

Table 10*Barriers Related to Limited Resources*

Statements	Strongly Disagree (%)	Disagree (%)	Don't know (%)	Agree (%)
I do not have enough time to read research	8	38	17	37
I do not have access to the books and journals I need	7	31	9	43
I cannot think of any topics that are worth researching	3	35	12	50
Teachers are not given time to do research as part of their weekly workload	7	42	15	36
I do not know enough how to access to digital libraries and web-based sources of research	4	1	12	83
I have not access to English books or journals I need	6	32	25	37

According to Crookes and Arakaki (1999), lack of time is one of the factors that hinders teacher research Engagement (TRE). Nevertheless, in the present study, lack of time is one of the barriers but not as dominant as the above-mentioned ones and it received 37% agreement. The fourth group of TRE challenges were categorized as demotivators under which the majority or minority of seven related barriers were examined. Among these hindering factors, the most dominant one was the statement that "there is not tangible benefit or reward for teachers to be research engaged" with 82% agreement (Table 11). The other considerable demotivative factors were "the lack of doing research by the other colleagues" 76%. "Not necessity of being research engaged for keeping the job and getting promotion" (57%). "The lack of institutions encouragement" (66%), and "the findings of researches are not important for EFL teachers to modify their probably less beneficial teaching procedures" took the fifth place among these demotivators with 55% agreement. Two other less dominant demotivators were "my job is to teach not to research" received (17%) and finally the lack of teachers' interest in doing research gained the least amount of agreement (6%).

Table 11*Demotivators for Doing Research*

Statements	Strongly Disagree (%)	Disagree (%)	Don't know (%)	Agree (%)
Most of my colleagues do not do research	3	5	16	76
I am not interested in doing research	32	47	15	6
There is no tangible benefit or reward for being research engaged	1	9	8	82
Reading research is not necessary for keeping my job and getting promotion	2	27	14	57
The finding of researches is not important for EFL teachers	4	17	24	55
My institution dose not encourage me to read research	3	12	19	66
My job is to teach, not to do research	5	43	35	17

The last topic to be examined in this section was the economic preventing matters for RE which consisted of three items. These three financial and economic factors obtained a high degree of concord among respondents. The item "teachers are paid only for teaching time" 77%, "There is no financial support to do research" 86%, and "there is no financial support from the educational system to do research" 76% (Table 12).

Table 12*The Economic Preventing Matters*

Statements	Strongly Disagree (%)	Disagree (%)	Don't know (%)	Agree (%)
Teachers are paid only for teaching time	4	10	9	77
There is no financial support to do research	1	2	11	86
There is no financial support from educational system to do research	13	1	10	76

Qualitative Results

The results of the interviews were in line with the finding of the questionnaires findings. In addition, the interviewees pointed to some more challenges and barriers related to TRE such as lack of time in EFL teachers' weekly timetable to teach their textbooks, considering research as a voluntary activity without any assignment for EFL teachers to be research engaged, lack of time allocation for RE in teachers' yearly workload and low status of research in educational system.

As to lack of time in EFL teachers' weekly timetable to teach their textbooks, one of the teachers stated:

I'm usually overloaded with weekly 24 hours of full-time teaching and 15 hours as part-time job at private institutes. Therefore, no time is left for doing research.

The other similar barriers reported by most teachers were the optional choice of doing research without any obligation to integrate research into the teaching job and unimportance of doing research at schools as a part of the teachers' tasks or duties. For example, one of the teachers reported:

In our educational system, research is not as important as teaching. Policy makers expect us to teach whatever is presented in textbooks with a fixed centralized methodology.

In sum, the qualitative results revealed that the teachers attributed the challenges and barriers to both personal and institutional factors highlighting the fact that research engagement requires and inner drive partly motivated by job requirements.

4. Discussion and Conclusion

The persuasive arguments in favor of the advantage of being research engaged by the teacher replete the literature but the reality remains that teacher research is a minority activity in ELT (Borg, 2009). A significant rationale for teacher research is its contribution to the teachers' professional development and practical value of educational research (Anvarudin & Pervin, 2013). Hargreav and Fullan (1992) suggest a model of teachers' development which consists of three approaches to teacher development (TD).

The first approach argues that those teachers who equipped themselves with appropriate knowledge and skills have more ability to provide better learning opportunities for their learners, the second approach of TD deals with self-understanding and the third approach in this model suggest that "the process and success of teacher development depends very much on the context in which it takes place"(1992, p.13). Both quantitative and qualitative results in the present study also confirm the important role of personal and institutional factors in research engagement.

Among these three approaches, the importance of knowledge in teachers' professional development was highlighted in the first one. Other scholars for example Shulman (1987) also emphasizes that teachers need to have not only content knowledge but also they need to possess pedagogical knowledge. From this perspective of teacher development, in order to inform and improve their practice, teachers are expected to utilize research-based practice. Other researchers (see Mansilla & Gardner, 1997) also emphasize the centrality of research knowledge.

Another profound rationale for being research engaged teacher is that teaching is a dynamic, progressing process and the teachers' initial/pre-service education is not enough for them to have a progressive and effective career in teaching. In other words, "the ongoing development of this knowledge relies on the continuous in-service education and teacher professional learning"(Santora et al. 2013, p.123). Reviewing such significant issues on teacher professional development maintaining teacher research as a basic and efficient means to obtain such needed development, and drawing on our findings, there are some places of argument in the context of the study.

As a place of concern, we can recall one of the findings of this study where the average of frequently research engagement by IEFL teachers is 16% and it confirms the Borg's (2003) citation that teacher research remains a minority activity. The second concerning issue is that most of the participants agree on the beneficial impact of the RE on their teaching careers or professional development and based on the finding the lack of interest to TRE was very little (6%) but this internal factor was not able to overcome some external demotivators for teachers. As another substantial place of concern, we would argue strongly that since a majority of respondents emphasized to be research engaged for getting a promotion or increasing their precedence to develop their status among their colleagues, this will make TR a non-collaborative activity with purposes other than teachers' knowledge and skills development.

The final remarkable part of the study was about the preventive factors or obstacles for teachers to be research engaged. From the five subtitles under which various obstacles were investigated economic matters obtain the highest mean (4.6) and the second important group of the barriers were non-collaborative school culture ones with the mean of (3.6). Consequently in line with the findings of the third part of the study (the reasons for being researched engaged), here it can be concluded that the most considerable barriers for teachers are external factors such as financial matters or non-collaborative school cultures.

The present study finding pertaining to Iranian EFL teachers' low level of engagement is in line with results reported by Mehrani and Behzadnia (2013). Economic constraints and lack of institutional support as dominant barriers are also consistent with Mehranirad and Behzadpoor (2022). However, the results reported by Li and Xu (2024) in a university context show higher engagement because of institutional support and role expectations.

Based on both quantitative and qualitative results, it can be concluded that limited research engagement among school-level EFL teachers is a global pattern caused by personal factors and

institutional barriers reinforced by economic and cultural constraints. In order to address these issues, systemic changes in the educational system are required and more longitudinal studies should be conducted on how to motivate teachers for more engagement in research. Teachers should be encouraged to use more resources in addition to books and to continue their education for more training on how to do research. The present study can have some implications for teachers, institutional administrators, and policy makers to seek professional development, foster collaborative culture among teachers through supporting research communities, and reduce teaching load to allocate time for research.

We acknowledge that although this study provided a considerable volume of quantitative and qualitative data from EFL teachers, our respondents represent a small nonrandom sample, and we cannot make any claim about the statistical generalizability of our findings. More studies can be conducted to consider other variables such as specific contexts (urban vs. rural and public vs. private schools), and the role of technology in overcoming research engagement barriers and challenges.

References

- Ahmed, A., & Pervin, N. (2013). English-language teachers' engagement with research: findings from Bangladesh. *Professional Development in Education*, 41(1), 21–39. <https://doi.org/10.1080/19415257.2013.861860>
- Allison, D. & Carey, J. (2007). What do university language teachers say about language teaching research? *TESL Canada Journal*, 24(2), 61–81. <https://doi.org/10.18806/tesl.v24i2.139>
- Allwright, D. & Bailey, K. M. (1991). *Focus on the language classroom: An introduction to classroom research for teachers*. Cambridge University Press.
- Allwright, D. (1997). Quality and sustainability in teacher research. *TESOL Quarterly*, 31(2), 368–370. <https://doi.org/10.2307/3588054>
- Atay, D (2008). Teacher research for professional development. *ELT Journal*, 62(2), 139–147. <https://doi.org/10.1093/elt/ccl053>
- Bai, L., & Millwater, J. (2011). Chinese TEFL academics' perceptions about research: An institutional case study. *Higher Education Research and Development*, 30(2), 233–246. <https://doi.org/10.1080/07294360.2010.512913>
- Borg, S. (2003). Teachers, researchers, and research in TESOL: Seeking productive relationships. *TESOL Matters*, 13(3), 1–4.
- Borg, S. (2009). English language teachers' conceptions of research. *Applied linguistics*, 30(3), 358–388. <https://doi.org/10.1093/applin/amp007>
- Borg, S. (2010). Language teacher research engagement. *Language Teaching*, 43(4), 391–429. <https://doi.org/10.1017/S0261444810000170>
- Borg, S., & Liu, Y. (2013). Chinese college English teachers' research engagement. *TESOL Quarterly*, 47(2), 270–299. <https://doi.org/10.1002/tesq.56>

- Bryman, A. (2007). Barriers to integrating quantitative and qualitative research. *Journal of Mixed Methods Research*, 1(1), 8-22. <https://doi.org/10.1177/2345678906290531>
- Burton, J. & Mickan, P. (1993). Teachers' classroom research: Rhetoric and reality. In J. Edge, J., & K. Richards, K. (Eds.), *Teachers develop teachers research* (pp. 113-121). Heinemann
- Cochran-Smith, M. & Lytle, S. L. (1999). The teacher research movement: A decade later. *Educational Researcher*, 28.7 15-25. <https://doi.org/10.2307/1176137>
- Creswell, J. (2003). *Research design: Qualitative, quantitative, and mixed-methods approaches* (2nd ed.). Sage.
- Crookes, G. & Arakaki, L. (1999). Teaching idea sources and work conditions in an ESL program. *TESOL Journal*, 8, 15-19. <https://doi.org/10.1002/j.1949-3533.1999.tb00151.x>
- Davies, P. (1999). What is evidence-based education? *British Journal of Educational Studies*, 47(2), 108-121. <https://doi.org/10.1111/1467-8527.00106>
- Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Dornyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research* (2nd ed.). Routledge.
- Everton, T., Galton, M., & Pell, T. (2002). Educational research and the teacher. *Research Papers in Education Policy and Practice*, 17(4), 373-401. <https://doi.org/10.1080/0267152022000031388>
- Freeman, D. (1998). *Doing teacher research*. Heinle and Heinle.
- Gao, X. & Chow, A. W. K. (2011). Primary school English teachers 'research engagement. *ELT Journal*, 66(2), 224-232. <https://doi.org/10.1093/elt/ccr046>
- Gu, L., & Wang, J. (2006). School-based research and professional learning: An innovative model to promote teacher professional development in China. *Teaching Education*, 17(1), 59-73. <https://doi.org/10.1080/10476210500528079>
- Hubbard, R.S., & Power, B.M. (1993). Finding and framing a research question. In L. Patterson, C. M. Santa, K. G. Short, & K. Smith (eds.), *Teachers are researchers: Reflection and action* (pp. 19-25). International Reading Association.
- Li, Y., & Xu, L. (2024). Exploring the influence of teachers' motivation, self-efficacy, and institutional support on their research engagement: A study of Chinese university EFL teachers. *System*, 121, 103272. <https://doi.org/10.1016/j.system.2024.103272>
- Mansilla, V.B. & Gardner, H. (1997). Of kinds of disciplines and kinds of understanding. *The phi delta kappan*, 78(5), 381-386.
- McDonough, J., & McDonough, S. (1990). What's the use of research? *ELT Journal*, 44(2), 102-109. <https://doi.org/10.1093/eltj/44.2.102>
- Mehrani, M. B. (2015). English teachers' research engagement: Level of engagement and motivation. *Iranian Journal of Language Teaching Research*, 3(1), 83-97.
- Mehrani, M., & Behzadnia, A. (2013). English teachers' research engagement: Current barriers and future strategies. *International Journal of Foreign Language Teaching and Research*, 1(4), 19-33.

- Mehranirad, M. (2023). English teachers' motivation for research engagement. *Applied Linguistics Inquiry*, 1(1), 67-73. <https://doi.org/10.22077/ali.2022.5574.1004>
- Mehranirad, M., & Behzadpoor, F. (2022). A survey of EFL teachers' research engagement. *Journal of Applied Linguistics and Applied Literature: Dynamics and Advances*, 10(1), 79-98. <https://doi.org/10.22049/jalda.2022.27469.1365>
- Nunan, D. (1992). *Research methods in language learning*. Cambridge University Press.
- Shulman, L.S., (1987). Knowledge and teaching: foundations of the new reform. *Harvard Educational Review*, 57(1), 1-22. <https://doi.org/10.17763/haer.57.1.j463w79r56455411>
- Sippel, L., & Sato, M. (2022). Teacher beliefs about second language research and researchers: Different roles at research-oriented universities. *The Modern Language Journal*, 106(3), 583-598. <https://doi.org/10.1111/modl.12797>
- Worrall, N. (2004). Trying to build a research culture in a school: Trying to find the right questions to ask. *Teacher development*, 8(2-3), 137-148. <https://doi.org/10.1080/136645304002000020241>
- You, Y. (2007). A deep reflection on the "key school system" in basic education in China. *Frontiers of Education in China*, 2(2), 229-39. <https://doi.org/10.1007/s11516-007-0019-6>

Appendix: English Questionnaire

➤ About Yourself:

Name : Gender:

1. Where do you teach English in most often (Tick ONE)

- a) Primary School
- b) High School
- c) Higher Secondary School
- d) Private language institutions

2. Your highest academic/professional degree:

Degree (e.g., BA or PhD):

Concentration/Major (e.g., ELT or literature)

3. Years of experience as an English teacher:

4. Type of institution where you teach:

- a) Private
- b) Public

➤ About Reading and Using Research Publication/s:

5. Have you ever done publishable research related to EFL teaching or learning?

- a) Yes
- b) No

6. How frequently do you read published language teaching research? (Tick ONE)

- a) Always b) Often c) Sometimes d) Rarely e) Never

7. Which of the following do you read?

- a) Books
b) Academic journals (print)
c) Academic journals (electronic)
d) Professional Magazines/Newsletters (Tick all that apply)

8. How frequently do you do research yourself? (Tick ONE)

- a) Always b) Often c) Sometimes d) Rarely e) Never

Reasons for being research engaged and the challenges toward TRE

Answer the following questionnaire (Tick one of the items: Strongly disagree, disagree, don't know, agree, strongly agree)

	Reasons for research engagement	Strongly disagree	disagree	I don't know	agree	Strongly agree
		1	2	3	4	5
1	reduces teachers' feelings of frustration and isolation					
2	creates a problem-solving mindset that helps teachers when they consider other classroom dilemmas					
3	allows teachers to become more reflective, critical, and analytical about their teaching behaviors in the					
4	to find better ways of teaching					
5	Because it allows me to collaborate with colleagues					
6	It causes to improve my teaching procedures					
7	helps empower teachers to influence their own profession at classroom, district, state and national levels					
8	allows teachers to move out of a submissive position and be curriculum innovators					
9	improves teachers' instructional decision-making processes					
10	Contributes to update my information about language teaching					
11	increases my professional status in teaching					
12	It helps me to solve my professional problems easily					
13	develops teachers' capacity for autonomous professional judgments					
14	To contribute to the improvement of my department/ institution					
15	Because it is good for my professional development					
16	I feel that doing research is an important part of my job					

	For research engagement Challenges and barriers	Strongly disagree	disagree	I don't know	agree	Strongly agree
		1	2	3	4	5
1	I face difficulty in understanding research reports/articles					
2	I do not know how to access to digital libraries and web-based sources of research					
3	I do not have access to the books and journals I need					
4	My institution (e.g., principal or department head) does not encourage me to read research					
5	Reading research is not necessary for keeping my job and getting promotion					
6	There is no financial support from educational system to do research					
7	I cannot think of any topics that are worth researching.					
8	My job is to teach, not to do research					
9	teachers are not given time to do research as part of their workload					
10	I have not access to English Language teaching books and journals					
11	I am not interested in doing research or using the research finding to improve my teaching					
12	Other teachers would not cooperate if I asked for their help					
13	Learners would not cooperate if I did research in class					
14	principal or department head) does not cooperate me to do research					
15	There is no financial support to do research					
16	There is no tangible benefit or reward for being research-engaged					
17	Teachers are paid only for teaching time.					
18	The findings of research are not important for EFL teachers					
19	I do not know enough about research methods					
20	Most of my colleagues do not do research					
21	My employer does not support					
22	. I do not have enough time to read research					
23	I think of research as an academic, large scale, statistical and technically difficult activity.					
24	Most research publications are not relevant to my teaching context					

Thank you for taking the time to respond to these questions. I appreciate your contribution to my research.



Contents lists available at [JSLP](#)

Journal of Second Language Pedagogy

Journal homepage: <https://www.sanad.iau.ir/journal/jslp>

The Effect of ChatGPT Implementation on Right- and Left-Hemisphere Dominant Iranian EFL Learners' Argumentative Writing and Metacognitive Awareness

Hassan Alizadeh Mahmoud Alilo

Department of English, Ta.C., Islamic Azad University, Tabriz, Iran

*Corresponding Author's Email: Hassan.alizadehmahmoudalilo@iau.ac.ir

KEY TERMS

Argumentative writing
ChatGPT
Left-hemisphere dominance
Metacognitive awareness
Right-hemisphere dominance

ABSTRACT

The current research investigated the impact of ChatGPT implementation on argumentative writing and metacognitive awareness among Iranian EFL learners, focusing on the moderating role of hemispheric dominance. It aimed to determine if ChatGPT differentially affects right- and left-hemisphere-dominant learners. A quasi-experimental design was employed with 60 intermediate participants aged 15-24 years from two institutes in Tabriz, Iran, who were allocated to experimental and control groups. Participants' English proficiency was measured using the Oxford Placement Test and their hemispheric dominance was determined as either left or right using the Open Hemispheric Brain Dominance Scale. Weir's rubric evaluated Argumentative writing, while metacognitive awareness was assessed using the Metacognitive Awareness Writing Questionnaire. ChatGPT was employed with the experimental groups in scheduled intervention for eight weeks; whereas, conventional instruction was given to the control groups. Data analysis through ANCOVA indicated that ChatGPT implementation significantly improved argumentative writing and metacognitive awareness among left-hemispherical learners compared to right-hemispherical learners. Small changes in control groups emphasized ChatGPT's effectiveness in analytical learners in particular. Pedagogically, this paper emphasizes the importance of personalized instruction by artificial intelligence (AI) to support cognitive variety and suggests incorporating AI in EFL to support analytical competencies tailored to learners' cognitive styles.

ARTICLE TYPE

Original Research Paper

Received:	3 March 2025
Revised:	3 June 2025
Accepted:	8 June 2025
Published Online:	1 July 2025

© The Author 2025

1. Introduction

The development of Artificial Intelligence (AI) has transformed EFL education and writing as a supportive element in language learning (Li et al., 2024). Argumentative writing, one important aspect of academic literacy, draws on thinking critically, connecting ideas coherently, and developing

metacognitive awareness (Farahian, 2015; Teng, 2025). Nevertheless, many EFL learners, including Iranian students, continuously encounter difficulties in acquiring such skills. Some features of AI-based tools like ChatGPT may support writing competency (Berk & Aydin, 2024; Ghafouri et al., 2024), but little is known about their unique effect on learners with different cognitive sets (Landrum et al., 2015), for example, with right- or left-hemisphere superiority. Since the association between the dominance of the brain hemispheres and specific language learning strategies has been established (Arabmofrad et al., 2021; Dülger, 2012), it is necessary to delve into the manifestation of ChatGPT in argumentative writing and the metacognitive awareness of Iranian EFL learners with different cognitive preferences.

Hemispheric dominance is crucial in determining how learners process information cognitively and linguistically (Suwanto & Hidayah, 2023). Individuals with left-hemisphere dominance are better at analytical, sequential, and logical reasoning, resulting in their superior ability in structured and rule-governed tasks, including tasks of syntax, grammar, and construction of logical arguments (Weisi & Khaksar, 2015). On the other hand, right-hemisphere-dominant individuals tend to lean towards holistic thinking, creativity, and context while processing information, which can contribute to their style of writing, coherence, and how well they can construct more nuanced arguments (Salehi et al., 2017). Therefore, being aware of learners' different cognitive styles may help determine how they react to writing help derived from AI. Although Esmaeil et al. (2023) and Mahapatra (2024) have focused on the general advantages of ChatGPT in writing lessons, the specific needs of left- and right-brain-dominant learners concerning AI tools have yet to receive adequate attention. Moreover, since metacognitive strategies are often employed and experienced differently by learners depending on their cognitive styles (Pitenoe et al., 2017), it becomes essential to explore how hemispheric dominance might shape learners' metacognitive engagement when interacting with AI-supported writing tools.

Metacognitive awareness, which controls the cognitive apparatus, is important in academic writing (Pitenoe et al., 2017). Metacognitive strategies such as planning, monitoring, and evaluating writing employed by effective writers make coherent and well-organized products (Shen & Tao, 2025). Moreover, studies have shown that students with high metacognitive awareness tend to perform better in writing tasks since they reflect on their cognitive processes (Teng, 2025). ChatGPT feedback is instantaneous and advises on structural issues, it could enhance metacognition (Abdelhalim, 2024). Nevertheless, whether right- or left-hemisphere dominant learners gain the same benefits from such AI-supported writing assistance is questionable. While the studies suggest the opposite, AI-written feedback potentially benefits analytical learners with strengths in logical structuring (Li et al., 2024), but creative learners may face challenges with its rigid, systematic nature (Jiang & Hyland, 2024). This difference highlights the importance of exploring the impact of ChatGPT implementation on metacognitive awareness concerning differing cognitive profiles (Tabib & Alrabeei, 2024).

Previous studies primarily focus on AI-supported writing instruction in general (e.g., Guo et al., 2024; Su et al., 2023; You et al., 2024) without accounting for the diversity in writing cognition among learners. Rai (2024) and Suwanto and Hidayah (2023) reveal that brain dominance affects language acquisition, as shown in the findings that right-brain-dominant students tend to have an abstract and intuitive style, while students with left-brain dominance prefer the analytical style (Ashraf et al., 2017; Lingard, 2023). This distinction is important in how AI-generated feedback affects people with different cognitive strengths and weaknesses.

Furthermore, Iranian EFL learners have special linguistic and cultural problems that can influence their writing. According to Moghadam and Jafarpour (2021), Persian is a right-to-left language with a dissimilar syntactic structure compared to English, which makes English structurally and conceptually difficult to learn. As a result, Persian-speaking students usually face difficulties regarding coherence, cohesion, and development in writing a coherent argument in English (Esfandiari & Allaf-Akbary, 2024). Also, different cultural beliefs toward writing and creativity could affect how much learners engage with AI-assisted instruction. Although ChatGPT streamlines the writing process (Zhang et al., 2025), its effectiveness for Iranian learners with diverse cognitive styles has not been widely studied.

Given that the Iranian EFL context lacks the implementation of the importance of critical thinking and self-regulation (Mohammadi et al., 2020), understanding how the integration of ChatGPT can play a role in such metacognitive awareness becomes a nation-specific concern. While studies have indicated that AI tools can help boost student engagement and motivation regarding writing tasks (Song & Song, 2023; Zare et al., 2025), ultimately, the effectiveness of AI tools will come down to how well they cater to learners' cognitive styles. If AI feedback primarily caters to left-brain learners, right-brain learners could struggle with the rigid nature of this feedback. On the one hand, if AI tools foster holistic thought and creativity, left-brain-dominant learners may flounder with impressionistic prompts. This potential discrepancy poses a nuanced inquiry into the efficacy of ChatGPT across cognitive styles.

Another critical aspect of this research is the comparative analysis of argumentative writing improvement between right- and left-hemisphere dominant learners. While previous studies have examined general improvements in writing skills with AI assistance (Bašić et al., 2023; Khampusaen, 2025), they have not specifically addressed cognitive differences. By examining whether ChatGPT disproportionately benefits one cognitive group over the other, this study will provide insights into the adaptability of AI-assisted writing instruction. Furthermore, investigating the impact on metacognitive awareness will reveal whether learners with different cognitive tendencies develop self-regulated learning strategies at comparable rates when using ChatGPT.

One major aspect of this research is the comparative examination of argumentative writing among right- and left-hemisphere dominant learners. Previous studies have explored general writing skills using AI assistance (e.g., Bašić et al., 2023; Khampusaen, 2025) but have not addressed cognitive differences. Analyzing the extent to which an AI tool such as ChatGPT gives undue advantage to one cognitive category or the other will afford us insights into the extent to which AI-supported writing instruction is an opportunity for innovation. Moreover, exploring its influence on metacognitive awareness will help identify if ChatGPT facilitates the development of self-regulated learning in learners with divergent cognitive inclinations at similar paces.

Without consideration of these cognitive differences, AI-assisted language learning can carry serious consequences. Whereas AI-based instruction primarily caters to sound analytical skills, this could lead to frustration and disengagement for creative learners (Urban et al., 2024). Conversely, when the tool fails to provide enough structure, left-brain-oriented learners may find it lacking coherence and logical flow in their writing (Khoiriah, 2019). It will be important to mitigate these potential disparities to provide equitable access to effective AI-assisted writing instruction. Moreover, a better understanding of the interaction between cognitive styles and AI tools can guide the development of individual learning strategies that will ultimately increase writing skills and metacognitive awareness.

Previous studies have increasingly explored the role of AI tools, particularly ChatGPT, in enhancing writing instruction within EFL contexts. Research has shown that AI-assisted writing platforms provide timely, structured, and grammar-oriented feedback that benefits learners' writing fluency and accuracy (Berk & Aydin, 2024; Ghafouri et al., 2024). For example, Esmaeil et al. (2023) and Mahapatra (2024) documented positive learner perceptions and improved writing quality when ChatGPT was integrated into writing instruction. However, these studies often treated learners as a homogeneous group, overlooking individual differences in cognitive processing styles. Meanwhile, studies examining brain hemispheric dominance have established its significant influence on language learning strategies and writing styles (Arabmofrad et al., 2021; Dülger, 2012; Suwanto & Hidayah, 2023). Learners with left-brain dominance typically prefer structured, analytical tasks, while right-brain-dominant learners excel in creative and intuitive reasoning (Salehi et al., 2017; Weisi & Khaksar, 2015). Despite this, little empirical work has connected hemispheric dominance with AI-assisted writing development. Moreover, although research supports the role of metacognitive awareness in writing proficiency—highlighting that learners who actively plan, monitor, and evaluate their writing tend to perform better (Farahian, 2015; Pitenoe et al., 2017; Teng, 2025)—few studies have investigated whether AI tools like ChatGPT influence metacognitive development differently based on cognitive style. Therefore, there remains a crucial gap in understanding how ChatGPT interacts with individual cognitive traits, particularly brain dominance, to influence both writing performance and metacognitive growth.

The study endeavors to fill the gap in the literature through an in-depth exploration of the influence of using ChatGPT on argumentative writing and the metacognitive awareness of right vs. left-hemisphere-dominant Iranian EFL learners. This study will address this gap in the field by investigating the interplay between AI-assisted feedback and cognitive styles, providing insight for a more holistic understanding of technology-enhanced language learning. Thus, these findings will enhance EFL pedagogy, AI-assisted education, and instructional design, leading to more inclusive and effective writing instruction. The growing use of AI in educational settings underscores the need to ensure that such tools support diverse cognitive profiles with equitable and personalized learning experiences for EFL learners and leads to the following research questions:

RQ1. Does using ChatGPT have any statistically significant impact on Iranian EFL learners' right and left hemisphere dominance in their argumentative writing?

RQ2. Does using ChatGPT have any statistically significant impact on Iranian EFL learners' right and left hemisphere dominance in their metacognitive awareness?

2. Methodology

The study followed a quantitative research approach, adopting a quasi-experimental design, which had a pre-test-post-test-control group design, requiring the existence of four groups: two experimental groups and two control groups. Quasi-experimental research is research that includes experimentation but is not truly experimental. Instead, their effects are based on the manipulation of the independent variable (Cook & Campbell, 1979). Participants are not randomly assigned to conditions or sequences of conditions. In this study, the experimental group was given written treatment by the innovative methodologies of ChatGPT, and the control group was instructed conventionally. Thus, the dependent variables in the scope of this study were argumentative writing

and metacognitive awareness, while ChatGPT implementation served as the independent variable. Moreover, the left hemisphere vs. right hemisphere dominance functioned as moderating variables.

2.1 Participants and Setting

The participants were selected from a population of 80 from two different institutes in Tabriz, namely Pardisan and Goldis. The final homogeneous sample consisted of 60 intermediate male EFL learners, aged 15–24 years, all native Persian speakers, recruited from the institutes following identical curricula. This method guarantees that participants have been exposed to a comparable academic background. According to the placement criteria of the institutes, they were intermediate students. Nonetheless, to guarantee the integrity and uniformity of the participants, a proficiency test was administered before the commencement of the primary research. From the proficiency test, selected candidates with scores in the range of 30–39 were classified as lower intermediate. Participants were then non-randomly allocated to the experimental or control groups. These included 30 students from Pardisan Institute in the experimental groups and 30 students from Goldis Institute in the control groups. Originally, there were 34 students in the two Goldis Institute classes, but to equalize their numbers with those of the Pardisan Institute, four participants were excluded. Additionally, participants completed a hemispheric dominance scale following the proficiency test to determine left-hemisphere vs. right-hemisphere dominant learners. All groups were then split into the left-hemisphere and right-hemisphere dominant groups. After administering the aforementioned scale, four groups were formed: two experimental groups, the Experimental Left-Hemisphere Dominant (ELHD) group and the Experimental Right-Hemisphere Dominant (ERHD) group, and two control groups, the Control Left-Hemisphere Dominant (CLHD) group and the Control Right-Hemisphere Dominant (CRHD) group, each consisting of 15 learners. The same instructor taught all groups to eliminate teacher confounding variables. The participants had already studied Evolve 1–3 level. They remained in Evolve 4 level for the duration of the study.

2.2 Instrumentation

To gather the data needed for the study, the researchers employed the following instruments at various stages of the study.

Oxford Placement Test (OPT)

The OPT developed by Dave (2004) was systematically applied to assess and verify whether the proficiency levels of the English language differed in any significant ways between the experimental and control groups investigated. The test is a well-structured formal evaluation divided into six levels of proficiency on the CEFR scale and assigns test scores to well-defined value boundaries for each of the discrete levels: Basic (A1: 0–17), Elementary (A2: 18–29), lower intermediate (B1: 30–39), upper intermediate (B2: 40–47), advanced (C1: 48–54), and very advanced (C2: 54–60). These categorizations conform to established standards of language proficiency, enabling the assessment of participants' skills. The OPT results collected at the onset of the study were vital in that they provided researchers with the ability to intentionally select individuals whose scores fell into the Lower Intermediate (B1: 30–39) range to maintain uniform language proficiency standards within the groups.

Open Hemispheric Brain Dominance (OHBD) Scale

The OHBD scale was created by Jorgenson (2015) to identify left-hemisphere vs. right-hemisphere dominance, used in the current research to split participants into right-brain and left-brain dominant groups. The questionnaire comprised 20 items on a five-point Likert-type scale from disagree (1) to agree (5). The scale ranges from 20 to 100, which is the minimum and maximum possible score. Per this questionnaire, the participants whose total score was between 20-55 were thought of as left hemisphere dominant, and learners whose total score was between 65-100 were treated as right hemisphere learners. The scale was translated into Persian, and the validity of the translated version was confirmed by two experts. To measure the reliability for internal consistency, Cronbach's alpha was used and showed a high coefficient of 0.89.

Argumentative Writing Tests

During the data collection procedure, participants were required to write two argumentative essays (pre-test and post-test). The pre-test was done one session before the treatment began, and the post-test was performed after the eight-session intervention was finished. The main purpose was to assess the written ability of students taking part in the intervention. The essays were scored using Weir's (1990, as cited in Ahour & Mukundan, 2009) analytic rubric, which is a validated tool designed to measure argumentative writing performance in important areas, such as ideas and content, organization, coherence and cohesion, voice, sentence fluidity, and grammatical accuracy. The writing topic was selected from NTC's TOEFL materials and was carefully adapted for lower-intermediate learners to ensure accessibility in terms of vocabulary, structure, and task demands. Although originally designed for higher proficiency levels, TOEFL-style prompts were used to expose learners to authentic academic tasks while maintaining an achievable level of complexity. The adapted version allowed participants to demonstrate key argumentative skills, such as stating a position, providing reasons, and organizing ideas coherently, within a manageable linguistic range. The prompt was reviewed by two EFL instructors to confirm its appropriateness for the learners' level and was then shared with all participants. This prompt formed the basis of an essay task that participants were to carry out before and after the intervention, wherein they wrote an essay of about 150-250 words. Fifty minutes was the time allotted for each session. Inter-rater reliability was used to maintain the validity of the assessment process. To do this, the essays were rated by two independent raters to confirm the consistency of the scoring. High agreement between raters indicated reliable scoring. In cases of substantial disagreement, discussions were conducted between the raters to resolve differences and achieve consensus. Weir's analytic scale includes several assessment factors: the relevance and adequacy of content, the organization of composition, cohesion, spelling, punctuation, adequacy of vocabulary for the intended purpose, and grammatical accuracy. The first three criteria are related to fluency, and the other criteria touch on writing accuracy.

Metacognitive Awareness Writing Questionnaire (MAWQ)

The MAWQ, originally developed by Schraw and Dennison (1994) as a general metacognitive awareness inventory and later adapted and validated for EFL writing by Farahian (2015), was used in this study to assess participants' metacognitive awareness in writing. The MAWQ, which is a 55-item questionnaire, evaluates two key components: Knowledge of Cognition, which includes declarative, procedural, and conditional knowledge, and Regulation of Cognition, which encompasses planning, audience consideration, monitoring, online strategies, revision, and evaluation. To examine changes in learners' metacognitive awareness, the questionnaire was administered in printed form both before and after

the treatment. To ensure accessibility for participants, the questionnaire was translated into Persian, and the translated version was verified by two experts for accuracy and clarity. The overall possible score on the measure ranged from 55 to 275. The original study by Farahian (2015) reported Cronbach's alpha reliability values ranging from 0.67 to 0.91 across different subscales, confirming its reliability as a metacognitive assessment tool. In this study, the overall Cronbach's alpha coefficient was calculated at 0.82, further demonstrating the internal consistency of the instrument.

2.3 Research Procedure

This study utilized a pre-test and post-test quasi-experimental design to explore the impact of ChatGPT on the argumentative writing and metacognitive awareness of the Iranian EFL learners with right- and left-hemisphere dominance. The procedure was grounded in rigorous, reliable, ethical principles, systematic steps were implemented chronologically from November 2024 to January 2025 at Pardisan and Goldis Language Institutes, Tabriz, Iran. The study was approved by the Ethical Committee of Pardisan and Goldis Language Institutes, and they acted as the gatekeepers to ensure that ethical research matters were addressed before the study. Recruitment was conducted during regular class sessions. The researcher described the purpose, procedures, and voluntary nature of the study and assured participants of anonymity and confidentiality as well as the right to withdraw from the study without consequence. Informed written consent was obtained from all subjects using printed forms that were signed, returned, and kept in a locked filing cabinet accessible to the researcher.

To establish baseline equivalence, two assessments were given. To ensure homogeneity amongst participants, the OPT was first conducted in a 60-minute session where all participants took part. Answer sheets were collected manually and scored by the researchers using the official scoring key. Only participants with a scores within the range of 30–39 (lower-intermediate, B1) were considered to maintain consistency. Out of an original pool of 80 male students, only 60 were kept, with 20 ruled out because their scores were outside of this range. Results were entered manually in a spreadsheet, with a colleague double-checking the scores for accuracy. Additionally, the OHBD scale was used to help classify participants as left- or right-hemisphere dominant. It was handed out on paper as part of a 30-minute session. The participants completed it individually, their responses were manually collected, scored, and categorized: 20–55 (left-hemisphere dominant) and 65–100 (right-hemisphere dominant). They were then non-randomly assigned to one of four groups of 15: ELHD, ERHD, CLHD, and CRHD. Experimental groups (n=30) from Pardisan Institute and control groups (n=30) from Goldis Institute ensured inter-institutional uniformity. To control for teacher-related variation, all groups were taught by the same instructor, fluent in Persian and English.

Prior to the treatment, baseline data on argumentative writing and metacognitive awareness were collected. First, the argumentative writing pre-test was administered in a 50-minute session. Participants wrote a 150–250-word essay on the prompt: "Do you agree or disagree with the statement: ChatGPT access must be limited to students?" This topic was deliberately selected for its immediate relevance to the participants, as all were EFL students who had encountered or were aware of AI in their educational experience. The topic was relatable and thought-provoking, encouraging critical engagement and allowing students to draw on their perspectives and experiences with AI tools in academic contexts. Essays were handwritten on standardized answer sheets, collected manually, and scored by two independent raters using Weir's (1990, as cited in Ahour & Mukundan, 2009) analytic rubric, focusing on content, organization, cohesion, vocabulary, and grammar. Scores were recorded manually on a scoring sheet, and inter-rater reliability was calculated, with discrepancies

resolved through discussion to ensure consistency. Then, MAWQ was administered in a 40-minute session. It was distributed in print and recorded in the spreadsheet. The consistent application of these validated instruments, used in similar contexts (e.g., Farahian, 2015; Ahour & Mukundan, 2009), ensured reliability and validity, as supported by a high inter-rater reliability coefficient for the writing pre-test scores ($r = .945, p < .001$).

Over eight weeks, once a week for 50 minutes, the experimental groups were instructed using ChatGPT, whereas the control groups were instructed by conventional instruction. Each session opened with a 15-minute mini-lecture on the elements of argumentative writing (e.g., essay formulation, counterarguments). Over the next few weeks, participants worked with ChatGPT via their smartphones, engaging in scaffolded writing activities built around structured prompts. These included tasks such as “List three arguments for and against limiting student ChatGPT access” (for idea generation), “Improve this paragraph’s coherence by using appropriate transitions” (for organization and cohesion), and “Revise this statement to make it clearer and more specific” (for clarity and argumentative strength). Participants were also asked to copy-paste their own writing into ChatGPT and request suggestions for improvement, such as grammar corrections or alternative vocabulary. They, then revised their drafts based on the AI-generated feedback and were encouraged to reflect on the changes they made and the reasons why they made changes. This iterative process was monitored by the instructor, who ensured that all learners actively used ChatGPT and received equitable support during sessions. Participants kept hand-written reflective journals (one entry per session of ~100 words) to prompts such as “In what way did the feedback offered by ChatGPT strengthen your argument?” Journals were collected and locked away every week.

Control groups used teacher-centered methodologies, mirroring experimental structure without AI tools. They would then draft, get instructor feedback related to grammar and organization, and go through peer reviews after a 15-minute lecture.” Metacognitive reflection was prompted through oral Socratic questioning (e.g., “Why is this evidence persuasive?”). Weekly handwritten drafts were collected for consistent feedback.

Under identical conditions to the pre-tests, after conducting treatment, post-intervention data were collected. All data were stored and anonymized using participant codes (e.g., ELHD-01). Writing tests were scored on separate sheets, with rater scores averaged after consensus. The same classroom conditions (e.g., lighting, seating) and timing (morning sessions) were maintained across institutes to minimize external variables. Data analysis was conducted in SPSS 27, with pre- and post-test scores compared using one-way Analysis of Covariance (ANCOVA), ensuring statistical rigor.

3. Data Analysis

In order to answer the posed research questions, some calculations, statistical routines, and results were produced, which will be explained in detail in this section. The details about descriptive statistics of groups regarding the study variables (argumentative writing and metacognitive awareness) are illustrated in Table 1.

Table 1*Group Statistics*

Tests	Groups	N	Mean	SD	SE Mean
Argumentative Writing Pre-Test	ELHD Group	15	19.66	1.447	.373
	ERHD Group	15	19.80	1.740	.449
Argumentative Writing Post-Test	ELHD Group	15	31.86	1.597	.412
	ERHD Group	15	30.06	1.830	.472
MAWQ Pre-Test	ELHD Group	15	125.86	2.587	.668
	ERHD Group	15	125.93	3.348	.864
MAWQ Post-Test	ELHD Group	15	231.53	3.270	.844
	ERHD Group	15	225.93	3.348	.864

As Table 1 demonstrates, the mean score of the argumentative writing pre-test for the ELHD group is 19.66 (SD = 1.447, SE = .373), and the ERHD group had a mean of 19.80 (SD = 1.740, SE = .449). Following the intervention, both groups showed significant improvement in argumentative writing, with the ELHD group having a mean score of 31.86 (SD = 1.597, SE = .412) and the ERHD group up to 30.06 (SD = 1.830, SE = .472). Concerning the metacognitive awareness, which was measured through the MAWQ, the means for the pre-test were equivalent, with the ELHD at 125.86 (SD= 2.587, SE= .668) and the ERHD group at 125.93 (SD = 3.348, SE = .864). Post-test scores showed significant increases after the test, with 231.53 (SD = 3.270, SE = .843) and the ERHD group scoring 225.93 (SD = 3.348, SE = .864), signifying significant improvements in metacognitive awareness for both groups, where the ELHD group scored slightly higher than the ERHD group in both post-tests. Additionally, the Pearson correlation coefficient was used to evaluate inter-rater reliability and compare the consistency between both raters. These analyses are outlined in Table 2.

As Table 2 displays, for the pretest scores, the inter-rater correlation was almost perfect for the control group, as $r = .945$ ($p < .001$), i.e., excellent scoring consistency. In the same way, the pretest scores of the experimental group exhibited identical reliability ($r = .945$, $p < .001$), indicating that raters consistently applied the scoring criteria between groups at baseline prior to intervention. Post-test scores showed even greater agreement, with correlations climbing to $r = .993$ ($p < .001$) for the control and experimental groups, respectively. Considering this high level of inter-rater agreement, one can conclude that both raters provided an equal level of accuracy and objectivity in assessing participants' performance in the pretest. Moreover, Levene's Test of Equality of Error Variances was administered to find out whether the error variance of the dependent variables was equal across the groups. Table 3 demonstrates these results.

Table 2*Inter-Rater Correlation for the Argumentative Writing Test Scores*

		Rater 1	Rater 2
Pretest of Control Groups (Rater 1)	Pearson Correlation	1	.945**
	Sig.(2-tailed)		.000
	N	30	30
Pretest of Experimental Groups (Rater 2)	Pearson Correlation	.945**	1
	Sig.(2-tailed)	.000	
	N	30	30
Posttest of Control Groups (Rater 1)	Pearson/Correlation	1	.993**
	Sig.(2-tailed)		.000
	N	30	30
Posttest of Experimental Groups (Rater 2)	Pearson/Correlation	.993**	1
	Sig.(2-tailed)	.000	
	N	30	30

** . Correlation/is/significant at the 0.01 level (2-tailed).

According to Table 3, the results showed that for the argumentative writing post-test, the test statistic was $F(3, 56) = 0.326$, with a significance value of .807, indicating that the null hypothesis of equal error variance across all the groups cannot be rejected, thus confirming homogeneity of variance. For the MAWQ post-test, the test statistic was $F(3, 56) = 2.704$, with a significance value of .054, which is marginally above the .05 threshold, suggesting that the assumption of equal error variance across groups is also reasonably met. These findings support the use of ANCOVA for analyzing both dependent variables, as the assumption of homogeneity of variance is generally upheld.

Table 3*Levene's Test of Equality of Error Variances for Argumentative Writing and MAWQ Post-test Scores*

<i>Scores</i>					
Tests		F	df1	df2	Sig.
Argumentative Writing		0326	3	56	.807
Posttest					
MAWQ Post-Test		2.704	3	56	.054

Tests the null hypothesis that the error variance of the dependent variable is equal across groups.

Results for the First Research Question

An ANCOVA was performed to determine whether ChatGPT implementation has any statistically significant impact on argumentative writing in terms of Iranian EFL learners' right- and left-hemisphere dominance. In this analysis, the post-test scores for argumentative writing served as the dependent variable, while the pre-test scores were used as a covariate to control for initial writing abilities. The independent variable, representing different group conditions (such as AI-assisted writing versus traditional methods), was used to capture variations in hemisphere dominance.

Table 4

ANCOVA Results: Effects of ChatGPT Usage on Argumentative Writing Test Scores

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Corrected Model	1044.045 ^a	4	261.011	100.997	.000	.880
Intercept	276.635	1	276.635	107.043	.000	.661
Writing pre-test	.128	1	.128	.049	.825	.001
group	1040.856	3	346.952	134.252	.000	.880
Error	142.139	55	2.584			
Total	44549.000	60				
Corrected Total	1186.183	59				

a. R Squared = .880 (Adjusted R Squared = .871)

The results, presented in Table 4, indicate that the overall corrected model, including the intercept, pre-test, and group, was statistically significant ($F(4, 55) = 100.997$, $p < .001$), explaining 88.0% of the variance in argumentative writing post-test scores ($R^2 = .880$, Adjusted $R^2 = .871$). The intercept was also statistically significant ($F(1, 55) = 107.043$, $p < .001$), with a partial eta squared of .661, indicating that the overall mean of the post-test scores, adjusted for the covariate, is significant. The argumentative writing pre-test scores, serving as a covariate, showed a marginally significant effect on post-test scores ($F(1, 55) = 0.49$, $p = .825$), with a partial eta squared of .001, suggesting that baseline writing scores have little influence on post-test outcomes. Most importantly, the group variable, representing the extent of ChatGPT usage, had a highly significant effect on argumentative writing post-test scores ($F(3, 55) = 134.252$, $p < .001$), with a partial eta squared of .880, indicating that 88.0% of the variance in post-test scores is explained by the extent of ChatGPT usage, demonstrating a strong and statistically significant impact on argumentative writing performance among groups. To further explore the differences among groups, pairwise comparisons were conducted using Bonferroni-adjusted significance levels.

Table 5*Pairwise Comparisons of ChatGPT Usage Effects on Argumentative Writing Test Scores*

(I) group	(J) group	MD (I-J)	SE	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
ELHD Group	ERHD Group	1.796*	.587	.021	.189	3.403
	CLHD Group	8.261*	.588	.000	6.652	9.869
	CRHD Group	9.879*	.589	.000	8.265	11.492
ERHD Group	ELHD Group	-1.796*	.587	.021	-3.403	-.189
	CLHD Group	6.465*	.587	.000	4.858	8.072
	CRHD Group	8.083*	.591	.000	6.464	9.701
CLHD Group	ELHD Group	-8.261*	.588	.000	-9.869	-6.652
	ERHD Group	-6.465*	.587	.000	-8.072	-4.858
	CRHD Group	1.618	.593	.051	-.004	3.240
CRHD Group	ELHD Group	-9.879*	.589	.000	-11.492	-8.265
	ERHD Group	-8.083*	.591	.000	-9.701	-6.464
	CLHD Group	-1.618	.593	.051	-3.240	.004

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

As Table 5 represents, the results indicate the following key findings: The ELHD group performed significantly better than the ERHD group (MD= 1.796, $p = .021$), CLHD group (MD= 8.261, $p < .001$), and CRHD group (MD= 9.879, $p < .001$). The ERHD group showed significant differences compared to the CLHD group (MD= 6.465, $p < .001$) and CRHD group (MD= 8.083, $p < .001$), with lower performance levels. The CLHD group had significantly lower scores than the ELHD group (MD= -8.261, $p < .001$) and the ERHD group (MD= -6.465, $p < .001$). However, the difference between CLHD and CRHD groups was not significant ($p = .051$). The CRHD group had the lowest performance, significantly differing from the ELHD (MD= -9.879, $p < .001$) and ERHD groups (MD= -8.083, $p < .001$), but not significantly different from the CLHD group ($p = .051$). These results suggest that ELHD learners had the highest post-test argumentative writing scores; whereas, CRHD learners had the lowest. The differences between groups were statistically significant in most cases, indicating that hemisphere dominance plays a critical role in writing performance.

Results for the Second Research Question

This study also investigated whether using ChatGPT has a statistically significant impact on the right and left hemisphere dominance in the metacognitive awareness of Iranian EFL learners. The results are summarized in Table 6.

Table 6

ANCOVA Results: Effects of ChatGPT Usage on MAWQ Scores

Source	Type III Sum of			F	Sig.	Partial Eta
	Squares	df	Mean Square			Squared
Corrected Model	17136.687 ^a	4	4284.172	2076.101	.000	.993
Intercept	254.637	1	254.637	123.397	.000	.692
MAWQ pretest	330.637	1	330.637	160.226	.000	.744
group	16488.012	3	5496.004	2663.352	.000	.993
Error	113.496	55	2.064			
Total	2716859.000	60				
Corrected Total	17250.183	59				

a. R Squared = .993 (Adjusted R Squared = .993)

The results from Table 6 indicate that the overall corrected model, including the intercept, pretest, and group, was statistically significant ($F(4, 55) = 2076.101, p < .001$), explaining 99.3% of the variance in MAWQ post-test scores ($R^2 = .993$, Adjusted $R^2 = .993$). The pretest scores significantly predicted post-test scores ($F(1, 55) = 160.226, p < .001$), with a partial eta squared of .744, indicating that 74.4% of the variance is explained by baseline metacognitive awareness. Most notably, the group variable, representing the extent of ChatGPT usage, had a highly significant effect on MAWQ post-test scores ($F(3, 55) = 2663.352, p < .001$), with a partial eta squared of .993, suggesting that 99.3% of the variance is explained by ChatGPT usage, demonstrating an extremely strong and statistically significant impact on hemispheric dominance in metacognitive awareness. To further clarify the differences between groups, pairwise comparisons were conducted and are presented in Table 7.

Table 7*Pairwise Comparisons of ChatGPT Usage Effects on MAWQ Scores*

(I) group	(J) group	MD (I-J)	SE	Sig. ^b	95% Confidence Interval for Difference ^b	
					Lower Bound	Upper Bound
ELHD Group	ERHD Group	5.660*	.525	.000	4.224	7.096
	CLHD Group	35.667*	.525	.000	34.231	37.102
	CRHD Group	35.921*	.526	.000	34.481	37.360
ERHD Group	ELHD Group	-5.660*	.525	.000	-7.096	-4.224
	CLHD Group	30.007*	.525	.000	28.571	31.443
	CRHD Group	30.261*	.526	.000	28.821	31.701
CLHD Group	ELHD Group	-35.667*	.525	.000	-37.102	-34.231
	ERHD Group	-30.007*	.525	.000	-31.443	-28.571
	CRHD Group	.254	.526	1.000	-1.185	1.694
CRHD Group	ELHD Group	-35.921*	.526	.000	-37.360	-34.481
	ERHD Group	-30.261*	.526	.000	-31.701	-28.821
	CLHD Group	-.254	.526	1.000	-1.694	1.185

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

b. Adjustment for multiple comparisons: Bonferroni.

As Table 7 illustrates, the results, adjusted for multiple comparisons using the Bonferroni correction, revealed significant mean differences between all pairs of groups ($p < .001$), except for the comparison between the CLHD group and the CRHD group ($p = 1.000$), which showed no significant difference ($MD = -.254$). Specifically, the ELHD group, which received experimental AI-based intervention and exhibited left-hemisphere dominance, scored significantly higher on the MAWQ post-test than the ERHD group, which received the same intervention but exhibited right-hemisphere dominance ($MD = 5.660$, $p < .001$). The ELHD group also scored significantly higher than the CLHD group ($MD = 35.667$, $p < .001$) and the CRHD group ($MD = 35.921$, $p < .001$), both of which served as control groups with no ChatGPT intervention. The ERHD group scored significantly higher than the CLHD group ($MD = 30.007$, $p < .001$) and the CRHD group ($MD = 30.261$, $p < .001$). Additionally, the CLHD group scored significantly higher than the CRHD group ($MD = 2.54$, $p < .001$). These findings indicate that the use of ChatGPT, particularly in the experimental groups (ELHD & ERHD), significantly enhances metacognitive awareness and differentiates levels of right and left hemisphere dominance compared to the control groups (CLHD & CRHD), with the ELHD group consistently showing the highest scores and the CRHD group the lowest, except for the non-significant difference between the CLHD and CRHD groups.

4. Discussion and Conclusion

The results of this research provide strong evidence for the differential effects of ChatGPT on argumentative writing and metacognitive awareness of Iranian EFL learners having different hemispheric dominance profiles, answering the two fundamental research questions put forth. For the first research question, which sought to find out whether ChatGPT has a statistically significant impact on argumentative writing conditioned on left- or right-hemisphere dominance, the results are robust. ANCOVA analysis showed that ChatGPT use explained 88% of the variance in post-test argumentative writing scores, which was significant, such that the ELHD group outperformed all other groups, including the ERHD group. Such a structured approach might be more beneficial for learners with left brain dominance who are more logical and analytical in their processing style (Weisi & Khaksar, 2015); they add correctness to their sequential levels of predictions and additions, and it is not uncommon for these students to have strict accuracy aspects and logical coherence in their output. By contrast, right-hemisphere dominant learners, who tend to thrive through holistic and creative thinking (Salehi et al., 2017), improved only modestly, falling behind their left-hemisphere peers, perhaps due to a misalignment between their cognitive styles and the AI's pattern-oriented advice (Jiang & Hyland, 2024). This finding, in line with Arabmofrad et al. (2021) and Dülger's (2012) studies, suggests that cognitive styles have an impact on language learning outcomes and expands upon this knowledge in the context of AI-supported writing.

In contrast, as demonstrated by the control groups, CLHD and CRHD scored significantly lower on their post-test as compared to their experimental counterparts, further emphasizing ChatGPT's effectiveness in enhancing writing skills, in concordance with findings from Bašić et al. (2023) and Khampusaen (2025). An unexpected finding, however, was a non-significant difference between the CLHD and CRHD groups ($p = .051$) despite the established cognitive differences between such profiles. The results might be due to conventional teaching's one-size-fits-all method, which may have failed to capitalize on the distinct advantages of each group, resulting in their performance even out. On the other hand, the small sample size may have had insufficient statistical power to detect subtle differences, which should be investigated further. The better performance of the ELHD group than the ERHD group also echoes with those of Li et al. (2024), who also discovered a higher efficacy of AI-based feedback for analytical learners, but the current research extended the findings by identifying hemispheric dominance as a key dimension in their analysis where it has not previously been considered.

Concerning the second research question, which examined ChatGPT's effect on metacognitive awareness based on hemispheric dominance, the results were more pronounced. The ANCOVA indicated that ChatGPT's right use accounted for 99.3% of the variance of MAWQ post-test scores, where the ELHD group performed better than both the ERHD group and both control groups again. This large effect size indicates a huge role ChatGPT plays in development, not just of writing production but also of improvement in learners' proficiency in planning, monitoring, and evaluating their processes of writing, confirming Abdelhalim (2024) and Teng (2025), who argued for AI's capacity to promote metacognitive development. The ELHD group's higher scores may be a function of their organization aligned with ChatGPT's structured prompts, which facilitate explicit reflection and revision and highly logical tasks (Ashraf et al., 2017). In contrast, the ERHD group's marginally lower gains might suggest a preference for intuitive strategies less emphasized by feedback provided by ChatGPT (Suwanto & Hidayah, 2023), but their improvement over the control groups illustrates the broader usability of the tool.

One unexpected observation was that there was no significant difference between the CLHD and CRHD groups ($p = 1.000$) when the experimental groups diverged significantly. From this perspective, metacognitive growth is potentially stimulated in these two cognitive styles when teachers deliver and provide feedback on their work and when peers review their output alike, confirming and diverging from Pitenoe et al. (2017), who argued that there is a need for explicit training on metacognitive skills. While Berk and Aydin (2024) confirmed the facilitative role of AI-powered tools in self-regulation, the extraordinary gains in MAWQ scores (e.g., ELHD: 125.86 to 231.53) among experimental groups further substantiates the strategy concerning hemispheric dominance that is uniquely attributed in the study to help the users better off. This difference in motivation and action between ChatGPT and a control group resonates with the findings of Song and Song (2023) that students receiving AI assistance in that study showed more motivation, thus likely aiding the metacognitive gains expected to be accrued in our ChatGPT condition.

The importance of these results comes from the insight they provide into how cognitive profiles mediate the effectiveness of AI-assisted learning, a dimension that has been underexplored in previous research (Guo et al., 2024; Su et al., 2023). Learners with left-hemisphere dominance seem to benefit from ChatGPT's analytical scaffolding, generating more coherent, logically structured arguments (Zhang et al., 2025). This implies that right-hemisphere dominant learners, though still benefiting from technical thinking, may have to put in extra work to utilize their creative strengths within AI frameworks, a challenge blurted in Urban et al. (2024). This differential impact highlights the brittle simplicity of the one-size-fits-all assumptions of many AI studies (e.g., Mahapatra, 2024) while reinforcing the necessity of contextually appropriate instructional design.

The implications of this study on theory, practice, and pedagogy are multifaceted. By introducing AI as an influencing factor (Rai, 2024; Suwanto & Hidayah, 2023) at a theoretical level, the present research extends the literature related to brain dominance and language acquisition, revealing that cognitive styles not only define conventional learning strategies but also have an evolving relationship with technological mediation. The findings theoretically call for incorporating ChatGPT into EFL writing instruction, particularly in traditional contexts such as Iran, in which rote learning is often favored by writing teachers to promote their students' success on tests rather than nurturing their critical thinking (Esfandiari & Allaf-Akbary, 2024). This approach serves to facilitate both writing skills and metacognitive awareness, which accommodates the longstanding difficulties in coherence and argumentation exhibited by Persian-speaking learners, as noted in Zare et al. (2025). For policy, these findings highlight the importance of educational stakeholders keeping cognitive diversity in mind while adopting the technology to ensure equitable benefits based on learners' profiles. In this way, this study extends Tabib and Alrabeei's (2024) research, who used AI to improve cognitive skills, by illustrating the hemispheric elements driving their effects.

In conclusion, this study demonstrates that ChatGPT significantly enhances argumentative writing and metacognitive awareness among Iranian EFL learners, with left-hemisphere dominant learners reaping the greatest benefits due to their alignment with the tool's structured feedback. Right-hemisphere dominant learners, while improved, show slightly lesser gains, highlighting the need for adaptive strategies to support diverse cognitive styles. These findings bridge a critical gap in AI-assisted language learning, offering actionable insights for educators and policymakers to optimize technology use in EFL contexts. By revealing the interplay between hemispheric dominance and AI efficacy, this research underscores the importance of personalized approaches, ensuring that tools like ChatGPT foster inclusive and effective learning experiences. Its impact resonates beyond Iran,

contributing to a global discourse on leveraging AI to enhance linguistic and cognitive development in education.

Overall, the results of this research show that ChatGPT has an impressive role in improving argumentative writing and metacognitive awareness of Iranian EFL students, especially those with left hemisphere dominance, who benefited the most as their characteristics fit well with the up-to-date structured feedback the tool provides. This capability is largely driven by advances in Natural Language Processing, which allow ChatGPT to analyze user input, understand linguistic structures, and generate coherent, context-sensitive feedback that aligns with academic writing conventions. While findings are evident for right-hemisphere dominant learners, they are somewhat hindered and bear evidence of requiring adaptation of principles to foster success in diverse cognitive learners. These findings help bridge a crucial gap in the literature on AI-assisted language learning, providing actionable insights for educators and policymakers to harness the use of technology in the EFL context. This research sheds light on the relationship between such hemisphere dominance and AI efficacy and shows the importance of personalized approaches, ensuring tools like ChatGPT promote learning experiences that are both user-friendly and successful. The implications are twofold, one that extends beyond Iran, as it feeds into an international conversation about using AI in education to further language and cognitive development.

While the study makes a valuable contribution, some limitations do exist. The lack of randomization in the group allocation to either of the two interventions comes with a potential selection bias, where participants from different institutes may differ on characteristics that were not measured, such as motivation or prior exposure to this type of technology. Although 60 learners were enough for a statistical analysis, it limits the generalizability, especially for intermediate Iranian EFL learners. Moreover, the study only focused on male learners, overlooking the fact that potential differences due to gender in the approaches to learning styles, technology use, or performance in writing could impact the findings. The use of self-reported hemispheric dominance through the OHBD scale (Jorgenson, 2015) further complicates matters, as behavior may not always reflect neurologic truths (Khoiriah, 2019). Furthermore, the eight-week intervention does not assess long-term effects, which was a noted limitation in similar short-term studies (You, et al., 2024). Third, the study's emphasis on argumentative writing does not represent the entire spectrum of writing genres, and ChatGPT may contribute to narrative or descriptive tasks, domains where the strengths of the right hemisphere may excel (Shen & Tao, 2025). This study was delimited to lower-intermediate male Iranian EFL learners aged 15–24, to maintain a homogenous sample in terms of language proficiency, gender, and educational context to minimize confounding variables and enhance internal validity.

These limitations open up opportunities for future research. Broader, randomized controlled studies in varied populations with EFL may help to validate these findings and improve generalizability. Employing neuroimaging or more robust cognitive assessments could sharpen the categorization of hemispheric dominance, which could mitigate any limitations from self-reporting (Arabmofrad et al., 2021). Longitudinal studies that investigate the long-term potential for ChatGPT to alter writing and metacognition would build upon Lingard (2023), who reported on the expanding role of AI in education. Exploring other genres of writing or employing adaptive, AI-enabled features for right-hemisphere learners (e.g., creativity prompts) may further level its advantages, as suggested by Jiang and Hyland (2024). Investigating cultural factors, like specifics of Persian grammatical structures or differences in attitudes toward AI, may also illuminate whether these outcomes are culturally specific (Esfandiari & Allaf-Akbary, 2024).

References

- Abdelhalim, S. M. (2024). Using ChatGPT to promote research competency: English as a foreign language undergraduates' perceptions and practices across varied metacognitive awareness levels. *Journal of Computer Assisted Learning*, 40(3), 1261-1275. <https://doi.org/10.1111/jcal.12948>
- Ahour, T., & Mukundan, J. (2009). Analytic assessment of writing: Diagnosing areas of strength and weakness in the writing of TESL undergraduate students. *Iranian Journal of Language Studies*, 3(2), 195-208.
- Arabmofrad, A., Badi, M., & Rajaei Pitehnoee, M. (2021). The relationship among elementary English as a foreign language learners' hemispheric dominance, metacognitive reading strategies preferences, and reading comprehension. *Reading & Writing Quarterly*, 37(5), 413-424. <https://doi.org/10.1080/10573569.2020.1846005>
- Ashraf, H., Branch, T., & Yazdi, M. T. (2017). Brain dominance quadrants and reflective teaching among ELT teachers: A relationship study. *International Journal of English Linguistics*, 7(2), 63-72. <http://dx.doi.org/10.5539/ijel.v7n2p63>
- Bašić, Ž., Banovac, A., Kružić, I., & Jerković, I. (2023). ChatGPT-3.5 as writing assistance in students' essays. *Humanities and Social Sciences Communications*, 10(1), 1-5. <https://doi.org/10.1057/s41599-023-02269-7>
- Berk, E. H., & Aydin, S. (2024). The impact of artificial intelligence-powered writing assistance systems on metacognitive writing strategies. Retrieved from <https://doi.org/10.21203/rs.3.rs-5383311/v2>
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-experimentation: Design & analysis issues for field settings*. Houghton Mifflin.
- Dave, A. (2004). *Oxford placement test 2: Test pack*. Oxford University Press.
- Dülger, O. (2012). Brain dominance and language learning strategy usage of Turkish EFL learners. *Cognitive Philology*, 5(2), 1-13. Retrieved from https://rosa.uniroma1.it/rosa03/cognitive_philology/article/view/10131
- Esfandiari, R., & Allaf-Akbary, O. (2024). The role of ChatGPT-based instruction and flipped language learning in metadiscourse use in EFL learners' argumentative writing and their perceptions of the two instructional methods. *Teaching English as a Second Language Quarterly (Formerly Journal of Teaching Language Skills)*, 43(3), 27-52. <https://doi.org/10.22099/tesl.2024.49975.3277>
- Esmail, A. A. A., Dzulkifli, D. N. A. K., Maakip, I., Matanluk, O. O., & Marshall, S. (2023). Understanding student perception regarding the use of ChatGPT in their argumentative writing: A qualitative inquiry. *Jurnal Komunikasi: Malaysian Journal of Communication*, 39(4), 150-165. <https://doi.org/10.17576/JKMJC-2023-3904-08>
- Farahian, M. (2015). Assessing EFL learners' writing metacognitive awareness. *Journal of Language and Linguistic Studies*, 11(2), 39-51.

- Ghafouri, M., Hassaskhah, J., & Mahdavi-Zafarghandi, A. (2024). From virtual assistant to writing mentor: Exploring the impact of a ChatGPT-based writing instruction protocol on EFL teachers' self-efficacy and learners' writing skill. *Language Teaching Research*, 13(6), 216–241. <https://doi.org/10.1177/13621688241239764>
- Guo, K., Li, Y., Li, Y., & Chu, S. K. W. (2024). Understanding EFL students' chatbot-assisted argumentative writing: An activity theory perspective. *Education and Information Technologies*, 29(1), 1–20. <https://doi.org/10.1007/s10639-023-12230-5>
- Jiang, F., & Hyland, K. (2024). Does ChatGPT argue like students? Bundles in argumentative essays. *Applied Linguistics*, 4(3), 37–52. <https://doi.org/10.1093/applin/amae052>
- Jorgenson, E. (2015). Development of the open hemispheric brain dominance scale 1.0. Available from: <https://openpsychometrics.org/tests/OHBDS/development/OHBDS1.0.pdf>
- Khampusaen, D. (2025). The impact of ChatGPT on academic writing skills and knowledge: An investigation of its use in argumentative essays. *LEARN Journal: Language Education and Acquisition Research Network*, 18(1), 963–988. <https://doi.org/10.70730/PGCQ9242>
- Khoiriah, N. (2019). *Relationship between left brain dominance and students' writing ability at Grade XI SMA Negeri 2 Padangsidempuan* [Unpublished doctoral dissertation, IAIN Padangsidempuan].
- Landrum, A. R., Eaves, B. S., & Shafto, P. (2015). Learning to trust and trusting to learn: A theoretical framework. *Trends in Cognitive Sciences*, 19(3), 109–111. <https://doi.org/10.1016/j.tics.2014.12.007>
- Li, H., Wang, Y., Luo, S., & Huang, C. (2024). The influence of GenAI on the effectiveness of argumentative writing in higher education: Evidence from a quasi-experimental study in China. *Journal of Asian Public Policy*, 17(2), 1–26. <https://doi.org/10.1080/17516234.2024.2363128>
- Lingard, L. (2023). Writing with ChatGPT: An illustration of its capacity, limitations & implications for academic writers. *Perspectives on Medical Education*, 12(1), 261–270. <https://doi.org/10.5334/pme.1072>
- Mahapatra, S. (2024). Impact of ChatGPT on ESL students' academic writing skills: A mixed methods intervention study. *Smart Learning Environments*, 11(1), 1–9. <https://doi.org/10.1186/s40561-024-00295-9>
- Moghadam, M., & Jafarpour, N. (2021). A survey of part of speech tagging of Latin and non-Latin script languages: A more vivid view on Persian. *Language Arts*, 6(1), 75–90. <https://doi.org/10.22046/LA.2021.05>
- Mohammadi, R. R., Saeidi, M., & Ahangari, S. (2020). Self-regulated learning instruction and the relationships among self-regulation, reading comprehension, and reading problem solving: PLS-SEM approach. *Cogent Education*, 7(1), 1–22. <https://doi.org/10.1080/2331186X.2020.1746105>
- Pitenoe, M. R., Modaberi, A., & Ardestani, E. M. (2017). The effect of cognitive and metacognitive writing strategies on the content of the Iranian intermediate EFL learners' writing. *Journal of Language Teaching and Research*, 8(3), 594–600. <http://dx.doi.org/10.17507/jltr.0803.19>

- Rai, K. (2024). *Brain hemispheric dominance, metacognitive awareness, and perceptual learning style preferences as correlates of Senior secondary school students' academic achievement in biology* [Doctoral dissertation, Tezpur University].
- Salehi, S., Kazemi, F., & Aslikhosh, A. (2017). Brain dominance and gender differences in writing performance of Iranian EFL learners. *Journal of Applied Linguistics and Language Research*, 4(3), 94-104.
- Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology*, 19, 460-475. <https://doi.org/10.1006/ceps.1994.1033>
- Shen, X., & Tao, Y. (2025). Metacognitive strategies, AI-based writing self-efficacy and writing anxiety in AI-assisted writing contexts: A structural equation modeling analysis. *International Journal of TESOL Studies*, 7(1), 70-78. <https://doi.org/10.58304/ijts.20250105>
- Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, 14(2), 126-143. <https://doi.org/10.3389/fpsyg.2023.1260843>
- Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, 57(3), 17-32. <https://doi.org/10.1016/j.asw.2023.100752>
- Suwarto, S., & Hidayah, A. (2023). The analysis of the brain dominance and language learning strategy used by university EFL learners. *Journal of General Education and Humanities*, 2(1), 79-90. <https://doi.org/10.58421/gehu.v2i1.64>
- Tabib, F. M., & Alrabeei, M. M. (2024). Can guided ChatGPT use enhance students' cognitive and metacognitive skills? In A. Al-Marzouqi & S. A. Salloum (Eds.), *Artificial intelligence in education: The power and dangers of ChatGPT in the classroom* (pp. 143-154). Springer Nature Switzerland.
- Teng, M. F. (2025). Metacognitive awareness and EFL learners' perceptions and experiences in utilising ChatGPT for writing feedback. *European Journal of Education*, 60(1), 128-141. <https://doi.org/10.1111/ejed.12811>
- Urban, M., Děchtěrenko, F., Lukavský, J., Hrabalová, V., Svacha, F., Brom, C., & Urban, K. (2024). ChatGPT improves creative problem-solving performance in university students: An experimental study. *Computers & Education*, 215(1), 15-31. <https://doi.org/10.1016/j.compedu.2024.105031>
- Weisi, H., & Khaksar, Z. (2015). The effect of hemispheric dominance on Iranian EFL learners' creativity in writing. *International Journal of English and Education*, 4(2), 383-397. <https://doi.org/10.1371/journal.pone.00715>
- You, Y. L., Lin, A., Han, S. J., Chu, C. C., & Dai, C. Y. (2024). Using ChatGPT to assist writers with planning writing tasks in an EFL classroom. In F. M. Editor & S. S. Editor (Eds.), *Trust and inclusion in AI-mediated education: Where human learning meets learning machines* (pp. 235-260). Springer Nature Switzerland.
- Zare, J., Al-Issa, A., & Ranjbaran Madiseh, F. (2025). Interacting with ChatGPT in essay writing: A study of L2 learners' task motivation. *ReCALL*, 2(1), 1-18. <https://doi.org/10.1017/S0958344025000035>

Zhang, R., Zou, D., & Cheng, G. (2025). ChatGPT affordance for logic learning strategies and its usefulness for developing knowledge and quality of logic in English argumentative writing. *System*, 128(1), 13–26. <https://doi.org/10.1016/j.system.2024.103561>



Contents lists available at [JSLP](#)

Journal of Second Language Pedagogy

Journal homepage: <https://www.sanad.iau.ir/journal/jslp>

A Narrative Review of Textbook Checklists: Insights for Language Pedagogy

Hamidreza Zahedi Gholezo^{1*}, Maryam Ghadamyari²

¹ Ministry of Education, Khorasan Razavi Province, Education Department, Iran

² Standard Language Academy

³*Corresponding Author's Email: hamidreza.zahedigholezo.2580@gmail.com

KEY TERMS

Language Education Materials
Textbook evaluation
Checklists
Internal evaluation
External evaluation

ARTICLE TYPE

Review Paper

Received:	17 February 2025
Revised:	10 March 2025
Accepted:	23 May 2025
Published Online:	1 July 2025

© The Authors 2025

ABSTRACT

Many studies have revealed that textbooks are the main reference for teachers and learners in learning. Thus, the analysis of textbooks can not only introduce opportunities but also inform educators of the threats. Textbook evaluation encompasses the assessment of a textbook's merits and demerits, enabling educators to select the most appropriate and effective materials to meet the needs of their users. One method to evaluate textbooks recommended by educational scholars is the implementation of a checklist. The present study focused on reviewing 27 studies that dealt with textbook evaluation checklists. It was shown that applying checklists can help researchers to have a better evaluation of textbooks. Most of the checklists have focused their attention on aspects such as physical, cultural, content, and skills. However, there seems to be a dearth of objectives as well as emotional aspects as elements of textbook evaluation. To enhance the selection of educational materials for language education to serve the diverse needs of learners, it is essential to refine and expand checklist criteria.

1. Introduction

The learning-teaching process is mainly centered on textbooks and they are considered as a tool in that process (Bahar & Zaman, 2012). Textbooks have a facilitating role in what is described and explained by teachers and materials (Jusuf, 2018). Moreover, most of the exams which are administered in schools and national levels must be based on textbooks; therefore, they play an essential role in syllabus design (Hutchinson & Torres, 1994; Khandaghi Khameneh & Hashamdar, 2021). Textbooks also play a supportive role for both teachers and learners (Cunningsworth, 1995). Textbooks are important in meeting learners' needs (Lathif, 2015) and in transferring knowledge and information (Cunningsworth, 1995). While textbooks have a direct relationship with the teaching

status, teachers must evaluate their strength and weaknesses throughout the teaching-learning process.

One of the early definitions presented for evaluation includes “the systematic attempt to gather information to make judgments or to pass decisions” (Lynch, 1996, p. 18). Evaluation can be unbiased and valid when it depends on reliable equipments (Mukundan & Nimehchisalem, 2012). The evaluation of textbooks is essential, as teachers must experience satisfaction in selecting an appropriate textbook (Udenwa & Okoye, 2023). Additionally, this evaluation offers insights into both the opportunities and challenges educators will face (Shahmohammadi, 2018).

There are different types of processes proposed for textbook evaluation including pre-use or predictive evaluation, in-use evaluation, and post-use evaluation (Ahour & Ahmadi, 2012; Ismael & Fahady, 2022; Tomlinson, 2003; Van Dat, 2022; Wuttisrisiriporn et al, 2020; Zhang, 2017). Among these, pre-use or predictive evaluation is popular and common (Littlejohn, 2022) which gives priority to selecting textbooks that approach the objectives of the course (Shahzad et al, 2021). In pre-use evaluation there should be an overlap of materials used in class and students’ needs (Bahar & Zaman, 2016). Solhi et al (2021) have maintained that in this phase, evaluators make an assumption to analyze the probable efficacy of the materials for the stakeholders (Dongxing, 2020; Soe, 2024; Ypsilanti & Karras, 2024). In-use evaluation, as another type of textbook evaluation, checks the utilized materials in the classroom, making teachers aware so they can modify themselves and their teaching methodologies (Yüksel et al, 2023). It examines the textbooks in a real classroom and investigates their effectiveness focusing on merits and demerits (Ma, 2020; Monib, et al., 2020; Pirzad & Abadikhah, 2022). As the third category, post-use or retrospective or reflective evaluation provides valuable data by measuring the real impact of using textbooks on learners (Bhutto et al, 2022; Nguyen, 2023). It follows a reflective approach to assess the quality of the textbooks that have been used previously (Caldas, et al., 2023; Dallasheh, 2024; Rokhsari, 2022). According to Gholami et al (2017), this kind of evaluation is more valuable due to determining the pros and cons of certain textbooks following their constant use (Alsan & Kucuktepe, 2023).

There are still other aspects and indexes to be considered in evaluating course books. One of the indexes is external that provides an overall image of the structural principles (Samoudi & Mohammadi, 2021) and mainly underlines the physical dimensions such as the front cover, lists of content, and preface (Dewangga & Ghazali, 2020; Mega, 2021). Internal evaluation, in contrast, provides exhaustive and in-depth information about the materials of the book (Nolaputri, 2023) and requires meticulous practice over the previous one (Mao, 2023). As indicated by Almatard (2019), internal evaluation tackles the extent to which the textbooks conform to what the author asserted (Tang & Zheng, 2018) and examines if opinions and claims made by external evaluation are authenticated or not (Papadaki & Karagianni, 2023). This type of evaluation explores aspects such as the explanation and exhibition of the four skills, the appropriacy of the materials given in each skill and adaptation to different learning styles, the organization and categorization of the materials, and the suitability of the text (Karim, 2020). This provides a deeper interpretation in terms of various and certain criteria (Al Fraidan, 2012). Employing external and internal evaluation, professionals and experts can make sense of what has been presented by authors and what has been presented by textbooks (Mao, 2023). Two other terms ‘macro’ and ‘micro’ have also been used instead of ‘external’ and ‘Internal’ in English Language Teaching (ELT) materials evaluation (Isaee et al, 2023).

Another category of evaluation whose implementation leads to policy data and documentation, decision making, gauging learners' attainments, and enhancing instructional materials and educational schemes has two options including formative and summative (Sadiq & Saalh, 2024). Formative evaluation attempts to determine the important dimensions of textbooks reflectively to manipulate while the teaching occurs (Jasim & Jasim, 2024) to ameliorate the textbooks (Al-Zeebaree, 2022; Giannarou, 2021; Le, 2014), to address the learners' needs (Mohamadi, 2013), to engage learners in educational decisions, and to provide information for curriculum development (Akef, 2015; Jafarigohar & Ghaderi, 2013). Saeed (2017) claims that formative evaluation continuously assesses the progress throughout the advancement process. Summative evaluation as the other extent, provides a comprehensive, conclusive assessment of the curriculum or textbook (Youssara & Houssna, 2024; Zohoorian, et al., 2018; Zohrabi, 2011). Besides these two types of evaluation, Richards (2001) recommended the term "illuminative" evaluation which provides "a deeper understanding of the processes of teaching and learning that occur in the program, without necessarily seeking to change the course in any way as a result" (Richards, 2001, p.289). It aims to identify various dimensions (Stec, 2014) and how they operate or are being carried out (Al-Masri, 2010; Jamshidi & Soori, 2013; Kashoob, 2018) to highlight and explain new issues that arise (Zernadji, 2017).

2. Methods in Textbook Evaluation

Experts and professionals employ three major types of methods in the process of textbook evaluation. The first one is impressionistic which deals with interpreting and scrutinizing textbooks based on general impressions (AbdelWahab, 2013; Chihab, et al, 2023). It entails understanding some features like organization, topics and issues, and layout and visuals (Camacho, 2024; Saud, 2023; Yolanda, 2023). It is employed to elicit information and knowledge that are subjective (Moazam & Jodai, 2014).

The second method entails using a checklist which is the opposite side of impressionistic method and deals with textbook evaluation objectively rather than subjectively. Checklist method is a standardized way in which some characteristics are marked based on a directive way (Sahin, 2020; Youssra & Houssna, 2024). Aprilia and Neisya (2023) remarked that in order to evaluate and assess a textbook, the term "checklist" which is an implementable and applicable way, is valuable (Aulia, 2019; Syahid et al, 2024). Ghoorchaei et al (2021) articulated that the checklist method presents a general framework and structure which encompasses categories that are understandable to participants.

The third method is in-depth method. It is concerned with intentional examination on particular features; for instance, unit or exercise (Nastiti, 2019; Pouranshirvani, 2017) and provides a good examination of those units (Solikhah, 2020). Karimi et al (2015) claims that in-depth method relates to analyzing and interpreting specific features of the book. One of the deficiencies of this method is that the unit which is selected might not be a good representative of the book contents (Monbec, 2020; Raj, 2018). According to McGrath (2002), in depth method goes beneath the publisher's and author's claims (Marcé, 2019) to see whether the materials are expected to achieve the proposed outcomes (Alkhaldi & Oshchepkova, 2018).

3. A Review of Checklist-Based Evaluation Studies Since 2014

This section seeks to review a selection of the studies that have employed checklists as their evaluation tool within the past decade.

In a recent study Shang (2024) attempted to focus on updating textbooks to improve the effectiveness of English pedagogy and learning. The author attempted to develop a checklist for evaluation of PEP “(the series of senior high school English teaching and learning book published by People’s Education Press in 2019)” to encourage Chinese teachers to promote learners’ communicative competences successfully and efficiently. The checklist included 32 items under two sections: physical indexes and course content. The second part contained five sub-sections: general, linguistic proficiency, cultural literacy, learning capacity, and supported materials.

In another study by Wang, et al. (2023) the developed checklist consisted of three headings. The first one was associated with the view that the evaluative criteria must be related to the context. In the second one, scholars gave priority to recognizing learners’ needs. The third one concentrated on tasks and activities which can led students to be autonomous. Researchers recommended the checklist has to be used for finding the most suitable resource book for exam readiness.

Özkan & Aşık (2023) focused on the integration of technology and in terms of technological potentials. They used the organization for Information and Communication Technology (ICT) tools that was suggested by Luo and Lie (2012). The checklist consisted of 4 main parts. The first part was concerned with incorporating technology in textbook content. The second part pertained to considering activities which dealt with technology and the extent to which the activities of the textbook support technology. The third part referred to the examination of the language skills which supported the ICT principles. The last part contained technical challenges.

Papadaki and Karagianni (2023) attempted to investigate one of the textbooks used in Greek from external and internal aspects of the process of textbook evaluation. In external evaluation the focus was on the introduction, cover, and table of specifications. In the internal evaluation the researchers analyzed the sequence of the four skills. Some included items entailed learners’ levels of proficiency, the objectives, and the physical aspects, provision of authentic inputs in different skills, intentional communication, and distinctive tasks based on students’ requirements and learning preferences.

Dassanayake (2022) created a foundational structure for assessing localized textbooks. The development itself involved an online survey and a literature review. The checklist included general attributes, lesson content (focusing on pronunciation, vocabulary and characters, grammar and syntax, and exercises), localization and culture-sensitivity, as well as layout and appearance one criterion which was added as previously mentioned was the cultural aspect.

Lopez-Medina (2021) created a checklist based on content and language integrated learning (CLIL). She established her checklist to compare four Cs in CLIL (content, communication, cognition and culture) with five standards in American Council on the Teaching of Foreign Languages (ACTFL)’s standards: communication, culture, connection, comparison and community. The items of the checklist comprise general criteria, linguistic criteria and four Cs of CLIL and as a final question about consistency of CLIL principles. Rating for each item was administered on a 4-point Likert scale. It was recommended to act as a reflection tool especially useful for novice teachers.

In another checklist development study, Munir et al (2021) applied the Research and Development (R & D) approach employing exploration, development, and validation. The criteria pertained to aims and approaches, design and organization, content, skills, vocabulary, exercises and activities, methodology, attractiveness of the textbook and physical make-up, teacher’s manual, and practical considerations.

Atar and Erdem (2020) endeavored to form a checklist based on social linguistics. They attempted to assess EFL textbooks in social context perspectives and proposed a model for the employment of the instrument. The concluding version of the checklist comprised of six items focusing on sociolinguistic points of view. Other aspects like language environments as well as intercultural awareness were included in it.

Two other researchers, Caner and Celik (2020), developed a checklist in order to analyze the elements of communicative competence. The elements that were chosen included conversational proficiency, practical language skills and tactical language skills. They recommended the researchers to use the developed scale to find and evaluate the principles of communicative competence within textbooks.

Jahan et al (2020) developed the checklist primarily based on communicative aspect and communicative language teaching (CLT) approach in Pakistani English textbooks. The checklist encompassed the physical factors of the book, goals and purposes, supplementary materials, and order in which language skills and grammar should be taught and developed. It also analyzed visuals based on learners' needs, supplementary materials, contextual and authentic properties of tasks, and engagement level of grammar activities based on CLT principles.

Nasri et al (2020) examined textbook of "Got it" series based on the socio-cultural aspect from teachers and supervisors' perspectives. In order to do that, the investigators applied a scale with 16 closed-ended items elicited and extracted from recommendation and suggestion from multiple sources. It consisted of 4-point Likert-scale for five categories including intercultural competence, cultural sensitivity, cultural expressions, goals and purposes, and teaching capacity.

Şahin (2020) explored an evaluative checklist for ELT (English Language Teaching) textbooks. In order to do this, the author reviewed the literature and analyzed the localized situation comprehensively. The whole checklist encompassed 4 divisions (design, content, passage and activities, visual context). The first domain included 3 sub-divisions related to outer and inner aspect of the course book as well the component of the textbook. The second dimension had 4 parts related to the overall content, matching of content of the textbook with goals and objectives, accommodation of language to learners' needs, and the degree of freedom given to the students, and also the presentation of content. The third and fourth dimensions contained items which related to the practice or activities and pictures or images of the textbook.

In China, Lei and Soontornwipast (2020) focused their checklist evaluation on Intercultural Communicative Competence (ICC). It focused on cultural perceiving and intercultural communication abilities. Their checklist which was predominantly formed through teachers' views was prepared on a 5-point Likert scale and it included two dimensions of overall quality having five sub-dimensions and ICC attributes with three sub-dimensions. Researchers recommended consolidating "general attributes" and "teaching and learning contents" to "overall quality". ICC attributes related to cultural facets and activities and exercises.

In turkey, Ahmet (2019) developed an exhaustive checklist to evaluate and analyze ELT textbooks. The checklist dealt with six parts, specifically, practical considerations, language content, design and organization, language skills, exercises and activities, and cultural considerations. The checklist was recommended for both pre-use and post-use evaluation. Practical consideration generally focused on external evaluation. In language content, the researcher considered components such as vocabulary and grammar. In language skills, the researcher focused on the approach of the

book about four skills. In exercises and activities, the researcher considered aspects such as learner engagement. The checklist also investigated cultural elements. The 5-point Likert-scale was recommended for pre-use evaluation in order to recognize and identify the textbook suitability and in-use or post-use evaluation to determine the textbook usefulness.

Zokaeieh et al (2019) developed a critical checklist to protect against potential hegemonic, ideological and manipulative intentions in global or commercial textbooks. The researchers carried out two main steps to conduct this research. Firstly, experts elicited 270 items based on the exhaustive review of the literature and after revisions 30 items were deleted. During the first process, the items were sorted according to their significance to key areas and domains within English programs. Secondly, a semi-structured interview was conducted in order to collect supplementary data and verify the lucidity and suitability of the items. The final version of the checklist included syllabus attributions, methodological barriers, content, listening and speaking, reading and writing, artwork and utilitarian attributes, vocabulary and grammar, pronunciation, learners' considerations, and teachers' considerations. To achieve a comprehensive view, the researchers recommend the checklist to be used beside other instruments.

Ounis (2019) evaluated the 7th grade basic education English language textbooks from EFL teacher's points of views in Tunisia. The researcher took the physical appearance, content, related task and exercises of the book into consideration to achieve an intelligible evaluation. The questionnaire with 42-items which was used in evaluation process was built up around three main segments, physical appearance, content, and practical factors. The second phase comprised of five sub-segments, reading comprehension, vocabulary, pronunciation, grammar, spelling, and conversational purposes. Participants had to answer the items in 4-point Likert scale.

In Thailand, Wuttisrisiriporn and Usaha (2019) provided a checklist based on the Common European Framework of Reference for Languages (CEFR) to enhance CLT approach. Having reconsidered ELT textbook evaluation studies from 2014 to 2017 to determine criteria in evaluating textbooks, they prepared a prototype checklist based on the national syllabus design. It was divided into eight sections which dealt with layout, design, and physical makeup, unit organization, content, topics, and language, language teaching methods and activities, four language skills, vocabulary, grammar, and accompanied/supplementary materials. They recommended their 5-point Likert scale for to pre-use and post-use evaluation to assess the effectiveness of textbooks.

Kashoob (2018) proposed a checklist to provide a method for assessing and evaluating ELT materials. A design-based research methodology was employed. The checklist was reviewed and tested by participants from six colleges of applied science in Oman and other ranked institutions. The checklist had three major parts, items related to the principles of acquiring a language such as learners' confidence, learning styles, personalization of knowledge, using authentic materials, paying attention to communicative aspect of the language, teaching methodology, needs identifications which was divided into learners' needs, teachers' needs, and institutional needs.

Aljouei and Alsuhaibani (2018) aimed to assess the effectiveness of 'Traveller' course books from teachers' points of views. They used a checklist which was combined of two primary sections of background knowledge with 11 items and the textbook assessment form with the characteristics such as configuration and concept, ease of access, properness, exercises, and competence. Işik (2018) attempted to develop an extensive checklist for evaluating materials in a three- step process. The extensive evaluation included two checklists, one related to impressionistic evaluation (external evaluation) and another was used for in-depth evaluation (internal evaluation). The external evaluation

checklist contained items related to face validity, learners, context, outlying questions, learner's book, assessment tools, workbook, and realia. This checklist was presented in a 5-point Likert scale. The internal checklist had the following aspects of approach, purpose, syllabus and curriculum, linguistics, teachers and learners, classroom discipline, teaching, content, culture, vocabulary, skills, configuration and integration, pre-exercise, body, further investigation, administration and assessment as well as software.

Ahmadi Safa and Karampour (2018) developed a checklist including sections such as layout and formatting, activities and exercises, content and language skills, supplementary materials and teacher's guide, language types, methodology, vocabulary, and grammar. The development included the adaptation of six previously constructed checklists. They recommended using the checklist to achieve the prospects of teachers and learners. The researchers claim that the cultural aspect is another added value to the previously developed checklists. Alharbi (2017) constructed an evaluation framework focusing on five aspects. The first one was physical appearance and textbook structure and aesthetics. The second aspect considered aims and purposes exploring user comprehension of objectives. The third aspect evaluated teaching methodology and practices while questioning the efficiency of activities and the textbook utility in various methodologies. The fourth aspect included items addressing skills and proficiencies. Lastly, the evaluation aspect focused on how effectively the textbook assessed students' progress.

Sundar (2017) aimed to generate a tool for evaluating textbooks. The researcher undertook three workshops regarding the development and creation of the checklist. In the first phase, the researcher and specialist panel attempted to identify the standards and criteria that could be taken into account in the process of evaluating a textbook. These experts recommended to work out on physical, visual, content coverage and style of presentations and also evaluation aspects. On the basis of the opinion and idea that were elicited, the checklist was developed. This tool considered the total key elements like physical, content, and visual dimensions and was recommended to be applied in the analysis of other textbooks.

Laabidi and Nfissi (2016) attempted to evaluate the practicality of the book, "visa to the world" according to seven major characteristics from teachers' point of views regarding design and layout, complementary resources, content, tasks, activities, cultural illustration, and language methodology. The aim of the study was to examine the linguistic knowledge and linguistic proficiency that learners achieved in order to communicate with native and non-native speakers in English. The questionnaire with 43 items in 4-point Likert-scale was created. In design and layout, the appeal of the book, table of specifications, organization of the heading and sub-headings as well as instructions and list of sources were taken into account. In the second part, some other resources were mentioned that could be beneficial to teacher. The third one encompassed items concerned with the appropriateness of content with learners' needs and levels, and topic attractiveness. The fourth- and fifth-parts dealt with the level of matching between tasks and activities of textbook with learners' enthusiasm, and students' engagement to communication. In cultural illustration, the cultural aspect of the book was analyzed. The final part focused on examining the language skills and sub-skills.

Nimehchisalem and Mukundan (2015) used Delphi method to validate their developed checklist. Professionals' comments and precise anticipations were sought. The ultimate version of the checklist was provided with two headings including general attributes and learning-teaching content. They recommended their checklist, as an independent medium, to be used by teachers to identify the positive and negative points of the selected textbooks.

Hussin et al (2015) sought to develop a new checklist according to the studies in the review of the literature. They attempted to identify which criteria to include in the evaluation checklist and to what degree the items of the checklist represent the internal reliability. They implemented changes to the checklists of AbdelWahab (2013), Mukundan et al. (2011), Miekley (2005), and Daoud and Celce-Murcia (1979). In simpler terms, they modified some items in these checklists to create and develop their own checklist for the purpose of evaluating the introduction of new vocabulary in ELT textbooks. A 4-point Likert-scale as an exhaustive checklist was proposed due to the coverage of key aspects related to the introduction of new vocabulary in ELT textbooks. They proposed that the checklist can be used for the purpose of evaluating the presentation of the new vocabulary items in textbooks.

Demir and Ertas (2014) attempted to propose a checklist for ELT course books based on past studies. Having mentioned some shortcomings of over 30 previous checklists, such as lack of practicality and clarity, the researchers stated that there was no criteria and indexes to organize and categorize different sections of those checklists. Therefore, they set four fundamental parts for their checklists including subjects and contents, skills and sub-skills, layout and physical make-up, and finally practical considerations. The researchers declared that the recommended checklist was not the outcome of a scale formulation study; rather, it offered a functional alternative for textbook evaluation by gathering elements from established checklists in the field. Karamifar et al (2014) developed a localized evaluation benchmark for EFL textbooks. The researchers tried to specify some benchmarks for international textbooks (TopNotch, the third edition of Interchange series, and Four Corners) used in institutions in Iran according to teachers and learners' perspectives. The teachers' questionnaire consisted of 13 parts mostly attributed to visual aspects of the textbooks, questions concerning the four skills, teaching techniques applied, cultural aspect, and items correlated with supplementary materials. The other questionnaire related to students' views. The learners' questionnaire included 10 parts focusing on outward appearance, skills and sub-skills, subjects and modules, goals of subjects. The two questionnaires were designed based on a 6-point Likert scale.

2.2 Instrumentation

4. Discussion and Conclusion

There are various influential elements and factors that affect learning. However, many studies have revealed that textbooks are a main reference for both teachers and learners in the process of learning as most of the activities that teachers design in their classrooms revolve around textbooks. Textbook evaluation refers to the process in which the merits and demerits of the textbook are identified, allowing teachers to provide the most appropriate and suitable book. As a result, selecting an appropriate textbook among many options is essential. There are several approaches to evaluate textbooks including internal/external, in-use/post-use, or summative/formative. In the same vein, miscellaneous methods can be applied and used to choose a suitable textbook, one of which is the checklist method. As Budiarsih (2022) mentioned, checklists establish a framework and methodological approach to evaluating and assessing textbooks.

In order to evaluate a textbook via checklist method, numerous criteria and aspects have been proposed and considered by researchers throughout the literature of textbook evaluation. Physical aspect is one of the aspects which includes details such as considering layout and appeal of the textbook. Another important aspect refers to content such as skills and sub-skills which the textbook encompasses. By considering this point, scholars recommend that a good textbook should attend to all four skills and also components. Objectives and purposes of the textbook is another element that

checklists have to assign items to. Textbooks that do not consider the goals and aims of the curriculum must be revised. The cultural and supplementary materials that course books provide are another necessary factor. The following table presents a summary of the related studies from 2024 to 2014. It entails the checklists' approaches taken by different scholars toward textbook evaluation in implementing checklists.

Table 1

Related Textbook Evaluation Checklist Studies since 2014

	Author	Date	Checklists' Approach
1	Shang	2024	Elements of communicative competence
2	Wang et al	2023	Context, learners' needs, and tasks and activities
3	Papadaki and Karagianni	2023	External and Internal aspect of the textbook
4	Özkan & Aşık	2023	Technological aspect
5	Dassanayake	2022	Localization and culture-sensitivity
6	Lopez-Medina	2021	Content and language integrated learning (CLIL)
7	Munir et al	2021	Research and Development
8	Lei & Soontornwipast	2020	Intercultural communicative competence (ICC)
9	Ahmadi Safa and Karampour	2018	Culture
10	Caner and Celik	2020	Elements of communicative competence
11	Atar and Erdem	2020	Social perspective
12	Jahan et al	2020	Communicative and CLT principles
13	Nasri et al	2020	Socio- cultural aspect
14	Şahin	2020	Localization and learner needs
15	Ahmed ACAR	2019	Pre-use and post-use
16	Wuttisrisiriporn & Usaha	2019	Communicative Language Teaching (CLT)
17	Zokaeieh et al	2019	Critical pedagogy
18	Ounis	2019	External and internal
19	Kashoob	2018	Practicality
20	Işık	2018	External and internal
21	Aljouei & Alsuhaibani	2018	Activities, skills, and appropriateness of the textbook
22	Sundar	2017	physical, visual, content coverage and style of presentations
23	Alharbi	2017	Internal and external
24	Laabidi & Nfissi	2016	Linguistic Knowledge and proficiency
25	Nimehchisalem and Mukundan	2015	Independent medium
26	Hussin et al	2015	Presentation of Vocabulary in ELT textbooks
	Demir and Ertas	2014	Functionality
27	Karamifar et al	2014	Localization

By considering the different fundamental elements studied by researchers, it can be concluded that some checklists have dealt with technological aspects (Özkan & Aşık, 2023), communicate principles (Caner & Celik, 2020; Jahan, et al, 2020; Lei & Soontornwipast, 2020; Shang, 2024; Wuttisrisiriporn & Usaha, 2019). Moreover, there have been checklists that focused on social perspectives (Atar & Erdem, 2020; Lei & Soontornwipast, 2020; Nasri et al, 2020). Other researchers' checklists have regarded both external and internal aspects of the course books (Alharbi, 2017; Işık, 2018; Papadaki & Karagianni, 2023), while another orientation that checklists may be developed based on might be content and language integration learning (Lopez-Medina, 2021). Yet, other checklists have analyzed textbooks from the perspective of culture (Ahmadi Safa & Karampour, 2020; Dassanayake, 2022). Some other criteria for developing textbook evaluation checklists include research and development (Munir et al, 2021), critical pedagogy (Zokaeieh et al, 2019), practicality (Kashoob, 2018), functionality (Demir & Ertas, 2014), and localization (Karamifar, et al., 2014; Sahin, 2020).

Based on the reviewed studies, there seems to be dearth of research and improvements in developing checklists that focus on critical thinking and problem-solving. Future checklists may focus on evaluating the presence of inquiry-based learning opportunities as well as real-world applications in the content presented inside language education textbooks. Another dimension which needs more attention is emotional and social learning. Future developers may develop items that can assess how textbooks contribute to emotional and social learning, stimulating abilities like empathy, collaboration, and self-awareness which are all critical in language learning. More importantly, while assessment is an integral part of the teaching process (Bozorova & Shaxnoza, 2023), checklists formed based on formative assessment need to be developed.

References

- AbdelWahab, M. M. (2013). Developing an English language textbook evaluative checklist. *IOSR Journal of Research & Method in Education*, 1(3), 55-70. <https://doi.org/10.9790/7388-0135570>
- Ahmadi Safa, M., Donyaee, S., Sohrabi, S., Farahani, M., Khassemy, D., & Saeedpanah, E. (2018). First grade high school English textbook evaluation: Prospect I. *Critical Studies in Texts & Programs of Human Sciences*, 8(18), 1-26. <https://doi.org/10.22111/ijals.2020.5647>
- Ahmet, A. C. A. R. (2019). Selecting and evaluating ELT textbooks. *Turkish Studies*, 14(2), 1-14. <https://doi.org/10.7827/TurkishStudies.14673>
- Ahour, T., & Ahmadi, E. (2012). Retrospective evaluation of textbook "Summit 2B" for its suitability for EFL undergraduate students. *Journal of Educational and Social Research*, 2(5), 195-202. <https://www.richtmann.org/journal/index.php/jesr/article/view/11950>
- Akef, H. D. (2015). Evaluating the English textbook "Iraq Opportunities" book 6 for the 2nd intermediate stage. *Alustath Journal for Human and Social Sciences*, 215(1), 105-132. <https://doi.org/10.36473/ujhss.v0i215.609>
- Al Fraidan, A. (2012). Evaluation of two ESP textbooks. *English Language Teaching*, 5(6), 43-47. <https://dx.doi.org/10.5539/elt.v5n6p43>
- AlHarbi, A. A. M. (2017). Evaluation Study for Secondary Stage EFL Textbook: EFL Teachers' Perspectives. *English Language Teaching*, 10(3), 26-39. <https://dx.doi.org/10.5539/elt.v10n3p26>

- Aljouei, K. F., & Alsuhaibani, Y. A. (2018). Evaluating" Traveller" English textbook series from Saudi secondary school EFL teachers' perspectives. *English Language Teaching*, 11(12), 138-153. <https://doi.org/10.5539/elt.v11n12p138>
- Alkhalidi, A. A., & Oshchepkova, T. (2018). An analysis of English language theories: A case study. *Advances in Language and Literary Studies*, 9(4), 227-236. <https://dx.doi.org/10.7575/aialc.all.v9n.4p.227>
- Al-Masri, N. (2010). *An Evaluation of the Reading Texts & Exercises in SB & WB of English for Palestine-Grade 9* (Doctoral dissertation, The Islamic University-Gaza).
- Almatard, A. (2019). *Developing an evaluation checklist for identity in ESOL textbooks*. Arizona State University.
- Al-Zeebaree, S. I. T. (2022). An evaluation of the local ELT coursebook in Kurdistan: Sunrise 7 based on the checklist of McDonough and Shaw 1993. *International Journal of Early Childhood Special Education*, 14(5). <https://doi.org/10.9756/INTJECSE/V14I5.1014>
- Aprilia, F., & Neisya. (2023). Evaluating an EFL textbook: A good fit, adequate fit, a poor fit or totally wrong? *Indonesian Research Journal in Education [IRJE]*, 7(2), 487- 499. <https://doi.org/10.22437/irje.v7i2.27610>
- Atar, C., & Erdem, C. (2020). A sociolinguistic perspective in the analysis of English textbooks: Development of a checklist. *Research in Pedagogy*, 10(2), 398-416. <https://doi.org/10.5937/IstrPed2002398A>
- Aulia, K. N. (2019). *An evaluation of English textbook for senior high school students grade x published by erlangga* (Master's thesis, Jakarta: FITK UIN Syarif Hidayatullah Jakarta).
- Bahar, H. B., & Zaman, B. (2013). Significance of text book evaluation for the purpose of effective second language acquisition. *ISOR Journal of Humanities and Social Science*, 12(5), 70-77. <https://doi.org/10.9790/0837-1257077>
- Bhutto, F. A., Syed, H., Rajput, A. S., Shah, S. T., & Chachar, Z. (2022). Analysis of alignment between secondary level English textbooks and national curriculum: A comparative study of public and private schools in Sindh, Pakistan. *Journal of Language and Literacy Education*, 18(1), n1.
- Bozorova, H., & Shaxnoza, R. (2023). The importance of assessment, and its impact in learning process. *Молодые ученые*, 1(6), 90-94. <https://doi.org/10.5281/zenodo.8011375>
- Budiarsih, L. (2022). *Textbook evaluation: Models of checklist methods*. In Proceedings International Conference on Teaching and Education (ICoTE) (3, (1), pp. 11-16). <https://dx.doi.org/10.26418/icote.v3i0.55815>
- Caldas, V. H., Ochoa, J., Morales, C., & Calle, A. (2023). Pilot testing of three English texts at the National University of Education. *Kronos-The Language Teaching Journal*, 4(2), 47-58. <https://doi.org/10.29166/kronos.v4i2.4755>
- Camacho, E. M. A. (2024). Presentation of Grammar in English Please 2: Evaluation of a Colombian Coursebook. *MEXTESOL Journal*, 48(2), n2.
- Caner, M., & Çelik, A. (2020). The communicative competence elements in the foreign language textbooks: A descriptive case study on Turkish and English textbooks. *IARTEM e-journal*, 12(1). <https://doi.org/10.21344/iartem.v12i1.727>
- Chihab, M., Benmhamed, M., & Belmekki, L. (2023). Textbook evaluation: Exploring the nature and frequency of critical thinking questions in English textbooks. *IOSR Journal of Humanities and Social Science*, 28(6), 01-08. <https://doi.org/10.9790/0837-2806050108>
- Cunningsworth, A. (1995). *Choosing your coursebook*. Hinemann.

- Dallasheh, W. (2024). Content analysis of the EFL textbook units according to Bloom's taxonomy: Between theory and practice. *Journal of Higher Education Theory & Practice*, 24(4). <https://doi.org/10.33423/jhetp.v24i4.6960>
- Dassanayake, N. (2022). Development of an evaluation checklist for localized Chinese language textbooks in Sri Lanka. *Lingua Cultura*, 16(2), 257-269. <https://doi.org/10.21512/lc.v16i2.8778>
- Demir, Y., & Ertas, A. (2014). A suggested eclectic checklist for ELT coursebook evaluation. *Reading*, 14(2), 243-252.
- Dewangga, I., & Ghazali, I. (2020). English as a foreign language textbooks evaluation for tenth-grade students. *Journal of English Language and Pedagogy*, 3(1), 1-12. <https://dx.doi.org/10.36597/jelp.v3i1.4863>
- Dongxing, Y. (2020). An evaluation of a Chinese language textbook: From students' perspective. *US-China Education Review*, 10(1), 35-44. <https://doi.org/10.17265/2161-623X/2020.01.004>
- Gholami, R., Noordin, N., & Rafik-Galea, S. (2017). A Thorough Scrutiny of ELT Textbook Evaluations: A Review Inquiry. *International Journal of Education and Literacy Studies*, 5(3), 82-91. <https://dx.doi.org/10.7575/aiac.ijels.v.5n.3p.82>
- Ghoorchaei, B., Derakhshan, A., & Ebrahimi, A. (2021). An Evaluation of English Textbook "Prospect 2": Teachers and Teacher Educators' Perceptions in the Spotlight. *Journal of Modern Research in English Language Studies*, 8(2), 90-99. <https://doi.org/10.30479/jmrels.2020.13156.1624>
- Giannarou, E. K. (2021). Evaluating the Textbook Used in the Greek Centre of Further Merchant Marine Education: School of Mechanics. *International Journal of English Language Studies*, 3(3), 35-40. <https://doi.org/10.32996/ijels.2021.3.3.5>
- Hussin, N. I. S. M., Nimehchisalem, V., & Kalajahi, S. A. R. (2015). Developing a checklist for evaluating the presentation of new vocabulary in ELT textbooks. *International Journal of Language Education and Applied Linguistics*, 2, 27-38. <https://doi.org/10.15282/ijleal.v2.461>
- Hutchinson, T., & Torres, E. (1994). The textbook as agent of change. *ELT Journal*, 48(4), 315-328. <https://doi.org/10.1093/elt/48.4.315>
- Isaee, H., Barjesteh, H., & Nasrollahi Mouziraji, A. (2023). Screening EFL teachers' perception on 'Prospect 1': The case of internal and external evaluation. *International Journal of Research in English Education*, 8(1), 71-91. <https://ijreeonline.com/article-1-729-en.html>
- Işık, A. (2018). ELT materials evaluation: A system and criteria. *Theory and Practice in Language Studies*, 8(7), 797-812. <https://dx.doi.org/10.17507/tpls.0807.11>
- Ismael, A. L. R. K., & Fahady, S. S. (2022). Evaluating the content of english curriculum of intermediate school from EFL teachers' point of view. *Alustath Journal for Human and Social Sciences*, 61(3), 516-530. <https://doi.org/10.36473/ujhss.v6i13.1612>
- Jafarigohar, M., & Ghaderi, E. (2013). Evaluation of two popular EFL coursebooks. *International journal of applied linguistics and English Literature*, 2(6), 194-201. <https://dx.doi.org/10.7575/aiac.ijalel.v.2n.6p.194>
- Jahan, K., Mukhtar, S., Yasmin, F., & Mushtaq, M. (2020). An evaluation of English textbook for grade-6: application of communicative approach. *International Journal of Psychosocial Rehabilitation*, 24(7), 10203-10223.
- Jamshidi, T., & Soori, A. (2013). Textbook evaluation for the students of speech therapy. *Advances in Language and Literary Studies*, 4(2), 159-164. <https://dx.doi.org/10.7575/aiac.alls.v.4n.2p.159>

- Jasim, H. H., & Jasim, R. R. (2024). Iraqi teachers' perspectives on evaluating English for Iraq textbook grade 4 in terms of EFL textbook standards. *Journal of language studies*, 8(12), 292-312. <https://doi.org/10.25130/Lang.8.12.17>
- Jusuf, H. (2018). The models of checklist method in evaluating ELT textbooks. *Al-Lisan: Jurnal Bahasa (e-Journal)*, 3(2), 17-35. <https://journal.iaingorontalo.ac.id/index.php/al/article/view/422>
- Karamifar, M., Barati, H., & Youhanaee, M. (2014). Local evaluation criteria for global textbooks: A case study from Iran. *Theory and Practice in Language Studies*, 4(5), 923. <https://doi.org/10.4304/tpls.4.5.923-930>
- Karim, S. (2020). *Investigating alignment between pedagogic policy and practice: An English language programme evaluation at secondary level in Pakistan* (Doctoral dissertation, University of Sheffield).
- Karimi, M., Kalani, E., & Shakouri, N. (2015). Textbook evaluation: A Conscientious look at American English File Series. *Elixir Literature*, 80, 31021-31027.
- Kashoob, M. (2018). *Evaluating ELT Materials with Specific Reference to Colleges of Applied Sciences General Foundation Programme in Oman: Towards a Viable checklist* (Doctoral dissertation, University of York).
- Khandaghi Khameneh, A., & Hashamdar, M. (2021). Iranian EFL teachers' perspective towards the high school English textbook, Vision3: An evaluation based on communicative approach. *International Journal of Research in English Education*, 6(3), 96-116. <https://doi.org/10.52547/ijree.6.3.96>
- Laabidi, H., & Nfissi, A. (2016). Fundamental criteria for effective textbook evaluation. *EFL Journal*, 1(2), 141-159. <https://dx.doi.org/10.21462/eflj.v1i2.13>
- Lathif, M. (2015). *An evaluation of English textbook for the eighth graders of junior high school. English Unpublished Thesis*. Surakarta, Universitas Yogyakarta.
- Lei, W., & Soontornwipast, K. (2020). Developing an evaluation checklist for English majors' textbooks in China: focus on intercultural communicative competence. *Arab World English Journal (AWEJ)*, 11. <https://dx.doi.org/10.24093/awej/vol11no3.6>
- Littlejohn, A. (2022). *The analysis and evaluation of language teaching materials. The Routledge handbook of materials development for language teaching*. Routledge. <https://doi.org/10.4324/b22783-23>
- Lopez-Medina, B. (2021). On the Development of a CLIL Textbook Evaluation Checklist: A Focus Group Study. *TESL-EJ*, 25(1), n1. <https://tesl-ej.org/pdf/ej97/a17.pdf>
- Lynch, B. K. (1996). *Language program evaluation: Theory and practice*. Cambridge university press.
- Ma, F. C. M. (2020). Analyzing a language-driven content-based language teaching textbook using sociocultural theory. *Indonesian JELT: Indonesian Journal of English Language Teaching*, 15(2), 169-181. <https://doi.org/10.25170/ijelt.v15i2.1848>
- Mao, M. (2023). Evaluation and Design of Textbook. *International Journal of New Developments in Education*, 5(12). <https://doi.org/10.25236/IJNDE.2023.051212>
- Marcé, P. (2019). *Language for Specific Purposes (LSP) Business Spanish Textbooks: A Content Analysis* (Doctoral dissertation, The University of Iowa). <https://doi.org/10.17077/etd.eusl-0o4c>
- McGrath, (2002). *Materials evaluation and design for language teaching*. Edinburgh University Press.
- Mega, N. D. (2021). *The evaluation of English textbooks based on Litz's theory* (Doctoral dissertation, UIN Raden Intan Lampung).

- Moazam, I., & Jodai, H. (2014). Textbook evaluation: a reflection on total English (intermediate). *Theory and Practice in Language Studies*, 4(11), 2402. <https://doi.org/10.4304/tpls.4.11.2402-2409>
- Mohamadi, Z. (2013). Program evaluation on general English course: A case study at Tabriz University. *Journal of Language Teaching and Research*, 4(6), 1285-1297. <https://doi.org/10.4304/jltr.4.6.1285-1297>
- Monbec, L. (2020). Scaffolding content in an online Content and Language Integrated Learning (CLIL) module. *International Journal of TESOL Studies*, 2(2), 157-173. <https://doi.org/10.46451/ijts.2020.09.12>
- Mukundan, J., & Nimehchisalem, V. (2012). Evaluative Criteria of an English Language Textbook Evaluation Checklist. *Journal of Language Teaching & Research*, 3(6). <https://doi.org/10.4304/jltr.3.6.1128-1134>
- Munir, S., Suzanne, N., & Yulnetri, Y. (2021). Developing criteria of an English textbook evaluation for Indonesian senior high school. *Ta'dib*, 24(2), 217-228. <https://doi.org/10.31958/jt.v24i2.4551>
- Nasri, M., Namaziandost, E., & Zariholhosseini, E. (2020). A Textbook Evaluation of Socio-Cultural Contexts in "Got It Series". *Journal of Linguistics and English Teaching Studies*, 2(1), 18-29. <https://doi.org/10.46870/lets.v2i1.146>
- Nastiti, S. (2019). *The comparative study of higher order thinking skills on the reading tasks written in foreign and local textbooks for grade 8* (doctoral dissertation, Universitas Negeri Semarang).
- Nguyen, H. T. (2023). An evaluation of two business English course books, business partner b1+ and business partner b2: students 'and Teachers' perspectives. *VNU Journal of Foreign Studies*, 39(6), 136-161. <https://doi.org/10.63023/2525-2445/jfs.ulis.5158>
- Nimehchisalem, V., & Mukundan, J. (2015). Refinement of the English Language Teaching Textbook Evaluation Checklist. *Pertanika Journal of Social Sciences & Humanities*, 23(4).
- Nolaputri, D. R. (2023). *An Evaluation on the Design and Content of a General English Textbook from the Students and Teachers' Perspectives*. Proceedings Series on Social Sciences & Humanities, 12, 194-202. <https://doi.org/10.30595/pssh.v12i.796>
- Ounis, T. (2019). Evaluating the 7 th grade Basic Education English Language Textbooks in use in Tunisia. *American Research Journal of Humanities Social science*, 2(2), 22-27.
- Özkan, T., & Aşık, A. (2023). Investigating Technology Integration into English Language Coursebooks: A Systematic Evaluation. *Journal of Learning and Teaching in Digital Age*, 8(2), 291-302. <https://doi.org/10.53850/joltida.1249220>
- Papadaki, A., & Karagianni, E. (2023). Internal and External Textbook Evaluation. *Eximia*, 10, 52-86. <https://eximiajournal.com/index.php/eximia/article/view/252>
- Pirzad, F., & Abadikhah, S. (2022). An evaluation of Iranian tenth-grade English textbook: with a focus on language skills, activities and teachers' perceptions. *Journal of English Language Teaching and Learning*, 14(29), 171-197. <https://doi.org/10.22034/elt.2022.50068.2475>
- Pouranshirvani, M. (2017). The external evaluation of new English textbook" Vision1" for tenth-grade students in Iranian high schools from teachers' perspectives. *Specialty journal of language studies and literature*, 1(2), 11-21.
- Raj, A. (2018). Elementary Teacher Education in Sikkim: Evaluation of In-house English Teaching/Learning Materials (Learning Resources). *The Criterion: An International Journal in English*, 9, 217-226.
- Richards, J. C. (2001). *Curriculum development in language teaching*: Cambridge university press.

- Rokhsari, S. (2022). The Perception of General English Books among the Iranian University EFL Instructors. *Journal of Applied Linguistics Studies*, 1(2), 13-27.
- Sadiq, B. J., & Saalh, S. M. (2024). Evaluation “Headway Academic Skills (Level 3) From Teachers and Students’ Perspectives. *International Journal of Academic Research in Progressive Education and Development*, 13(1). <https://doi.org/10.6007/IJARPED/v13-i1/20726>
- Saeed, K. A. (2017). *Teachers’ Evaluation and Perceptions on Adaptation of Sunrise Coursebook at Sulaymaniyah Governorate in Kurdistan Region of Iraq* (Master's thesis, Eastern Mediterranean University (EMU)-Doğu Akdeniz Üniversitesi (DAÜ)).
- Şahin, S. (2020). Developing a checklist for English language teaching course book analysis. *International Journal of Education and Research*, 8(1), 107-120.
- Samoudi, N., & Mohammadi, M. (2021). An ELT textbook evaluation: A two-phase criterion. *Journal of English Language and Pedagogy*, 4(1), 1-8. <http://dx.doi.org/10.36597/jelp.v4i1.9666>
- Saud, W. I. (2023). Evaluation of Skills for Success-Reading and Writing Based on Criteria Designed by Alan Cunningsworth. *Technium Soc. Sci. J.*, 46, 331. <https://doi.org/10.47577/tssj.v46i1.9336>
- Shahmohammadi, S. (2018). Textbook evaluation: looking at Prospect series through teachers’ perspective. *Research in English Language Pedagogy*, 6(2), 182-204. <https://doi.org/10.30486/relp.2018.542578>
- Shahzad, A. K., Saeed, I. M., & Imtiaz, A. (2021). Evaluation of the efficacy of English textbook for 8th grade by Punjab Text Book Board. *Ilkogretim Online- Elementary Education Online*, 19(04), 6172-6186. <https://doi.org/10.17051/ilkonline.2020.04.765025>
- Shang, J. (2024). *Evaluating senior high school English textbooks in China* (Master’s thesis), University of Wisconsin-Platteville.
- Soe, T. (2024). Investigating Critical Thinking in ELT Textbooks: A Systematic Literature Review of Textbook Evaluation Studies. *Theory and Practice of Second Language Acquisition*, 1(10), 1-29. <https://doi.org/10.31261/TAPSLA.13882>
- Solhi, M., Mert, Y. Z., Çelen, Z., & Kısa, R. (2021). Evaluating the long-term effectiveness of English language coursebooks at Turkish public high schools. *Eurasian Journal of Applied Linguistics*, 7(1), 458-483. <https://dx.doi.org/10.32601/ejal.911493>
- Solikhah, I. (2020). Evaluating EAP Textbooks for Indonesian University Students Using Impressionistic and In-Depth Assessment. *Indonesian Journal of English Language Teaching and Applied Linguistics*, 4(2), 395-411. <https://dx.doi.org/10.21093/ijeltal.v4i2.606>
- Stec, M. (2014). Evaluation of ELT materials for young learners: Course-books as cultural artifacts. *Linguistics, Culture and Identity in Foreign Language Education*, 5, 1187-1196.
- Sundar, S. (2017). Development of Tool for the Evaluation of a Textbook. *Universal Research Reports*, 4(13), 495-503. <https://urr.shodhsagar.com/index.php/j/article/view/1169>
- Syahid, A. ., Arzaqi, D. A. ., Nuraisyah, D. ., Ramadhani, S. R. ., Sari, N., Al-Habsyi, M. N. ., & Nadia, A. A. (2024). Analyzing language proficiency for the clt principle: an evaluation of an eightgrade english textbook. *Jurnal Ilmiah Multidisiplin*, 3(02), 43-48. <https://doi.org/10.56127/jukim.v3i02.1278>
- Tang, X., & Zheng, D. (2018, October). *A retrospective evaluation of college ELT textbooks in China. In 2018 International Conference on Social Science and Education Reform (ICSSER 2018)* (pp. 57-61). Atlantis Press.
- Tomlinson, B. (2003). *Developing Materials for Language Teaching*. Continuum.

- Udenwa, V. C. U., & Okoye, C. C. O. (2023). Evaluation of English language textbooks used in the junior secondary schools in awka education zone, anambra state. *Journal of Theoretical and Empirical Studies in Education*, 8(1), 220-237.
- Van Dat, N. (2022). A Review on Approaches of Evaluating Tertiary Textbooks. *AJIR Proceedings*, 179-184. <https://doi.org/10.21467/proceedings.132.20>
- Wang, S., Li, C., & Yu, W. K. (2023). Assessing the quality of reference books on the English writing examination in Hong Kong: How can an evaluation checklist help?. *TESOL Journal*, 14(2), e710. <https://doi.org/10.1002/tesj.710>
- Wuttisrisiriporn, N., & Usaha, S. (2019). The Development of a Localized ELT Textbook Evaluation Checklist: A Case Study from Thailand. *Thaitesol Journal*, 32(2), 46-64. <https://so05.tci-thaijo.org/index.php/thaitesoljournal/article/view/229325>
- Wuttisrisiriporn, N., Vinitchevit, N., & Usaha, S. (2020). A Situational Analysis of EFL Textbook Selection in Thai Public Schools. *Indonesian TESOL Journal*, 2 (2), 71-80. <https://doi.org/10.24256/itj.v2i2.1247>
- Yolanda, M. (2023). *An analysis of writing task in "bahasa inggris when English rings a bell" textbook for seventh-grade students of junior high school* (doctoral dissertation, Uin raden Intan Lampung).
- Youssra, Z. E. G. H. M. A. R., & Houssna, B. O. U. D. J. E. M. L. I. N. E. (2024). *Evaluating the Effectiveness and Suitability of Third-Year Primary School Textbook Pronunciation Content* (Doctoral dissertation, university center of abdalhafid boussouf-MILA).
- Ypsilanti, A., & Karras, I. (2024). "Coursebook or No Coursebook"? A False Dilemma. *European Journal of Education and Pedagogy*, 5(2), 45-51. <https://doi.org/10.24018/ejedu.2024.5.2.808>
- Yüksel, M. E., Köse, R., & Doğan, Ç. (2023). An EFL coursebook evaluation through the lens of preparatory school instructors. *ELT Research Journal*, 12(1), 1-28.
- Zernadji, K. (2017). *A Pragmatic Evaluation of Speech Acts in My Book of English First Year Middle School Textbook* (Doctoral dissertation), University of Biskra.
- Zhang, X. (2017). A critical review of literature on English language teaching textbook evaluation: What systemic functional linguistics can offer. *Journal of Language and Cultural Education*, 5(1), 78-102. <https://doi.org/10.1515/jolace-2017-0005>
- Zohoorian, Z., MatinSadr, N., & Shamabadi, F. (2018). A Summative Evaluation of 'Prospect 1': Employing the ARCS Model. *International Journal of Instruction*, 11(3), 449-462. <https://doi.org/10.12973/iji.2018.11331a>.
- Zohrabi, M. (2011). Coursebook Development and Evaluation for English for General Purposes Course. *English Language Teaching*, 4(2), 213-222. <https://doi.org/10.5539/elt.v4n2p213>
- Zokaeieh, S., Najafi Karimi, Sh., Nouri, F., & Afsana, H. (2019). Preparing a critical questionnaire for evaluating language teaching books. *Journal of English Language Pedagogy and Practice*, 12(1), 202-225. <https://doi.org/10.30495/jal.2019.671933>



Contents lists available at [JSLP](https://www.sanad.iau.ir/journal/jslp)

Journal of Second Language Pedagogy

Journal homepage: <https://www.sanad.iau.ir/journal/jslp>

A Feasibility Study of Lesson Study Implementation for Iranian English Teachers' Professional Development: Managers' Insights

Reza Hadinezhad¹, Seyed Mohammad Reza Amirian^{2*}, Mohammad Davoudi³

¹ MA graduate, Department of English Language and Literature, Faculty of Literature and Humanities, Hakim Sabzevari University, Iran

² Associate professor, Department of English Language and Literature, Faculty of Literature and Humanities, Hakim Sabzevari University, Iran

³ Associate professor, Department of English Language and Literature, Faculty of Literature and Humanities, Hakim Sabzevari University, Iran

*Corresponding Author's Email: Sm.amirian@hsu.ac.ir

KEY TERMS

Collaborative learning
EFL teachers
Lesson study
Teacher professional development
Teacher reflection

ABSTRACT

Recent studies have identified Lesson Study (LS) as an effective professional development activity for language teachers, emphasizing its potential to enhance teaching practices through collaborative planning, observation, and reflection. This study investigated the feasibility of implementing LS for Iranian English teachers' professional development, using content analysis and quantitative analysis of code frequencies to assess LS's recognition and applicability. Semi-structured interviews were conducted with six institute managers, utilizing 13 open-ended questions. After transcribing the interviews, 14 codes were identified through deductive descriptive coding: belief, cognition, efficiency, implementation, preference, obstacles, financial issues, voluntary implementation, permission, class observation, interaction, collaboration, teacher, and teaching. Among these, cognition, efficiency, and implementation emerged as the most critical factors. Findings revealed that LS was not recognized by most English institute managers. However, its positive effects on EFL teachers' perceptions highlighted its value as a professional development approach. Despite these advantages, LS was deemed inapplicable in Iran due to a lack of relevant knowledge and understanding among educators. The study concludes that implementing LS in Iranian English institutes requires cultural and contextual adaptation. To establish LS as a professional development tool, it must be effectively promoted and understood within the local educational framework.

ARTICLE TYPE

Original Research Paper

Received:	4 March 2025
Revised:	3 June 2025
Accepted:	8 June 2025
Published Online:	1 July 2025

© The Authors 2025

1. Introduction

Lesson Study (LS) originated in Japan as a joint professional development approach for teachers. By using LS, teachers work together to plan lessons, observe classroom teaching, and engage in reflective discussions to enhance instructional practices. LS has become globally popular as an effective

approach for improving teacher learning and enhancing educational outcomes (Godfrey et al., 2019). In spite of its widespread success, one of the continuous challenges in professional development is the isolation that many teachers face, with restricted opportunities to observe peers or receive constructive feedback on their teaching methods (Coenders & Verhoef, 2019). LS, which has been used for over a century in Japan, has become a cornerstone of teacher growth (Schipper et al., 2017).

However, the use of LS in other educational contexts, particularly in non-Japanese settings, has been less extensively studied. Although professional collaboration is increasingly emphasized in education (Richit & Da Ponte, 2021), countries like Iran still face substantial obstacles to applying collaborative professional development models. Many ELT teachers in Iran continue to work in environments characterized by isolation and limited access to peer feedback. This separation raises critical questions about the feasibility and applicability of LS in such contexts. Although LS has been shown to foster reform-based teaching practices and improve teacher reflection (Arslan, 2019), its potential within the Iranian educational landscape remains largely unexplored.

The existing literature on LS, while extensive in its analysis of Western and East Asian contexts, provides little insight into how this model can be adapted for countries like Iran, where systemic and cultural differences may impede its success. There has been research by Dudley (2013) suggesting that LS can be successfully used in small groups of teachers, even with minimal resources, but these kinds of findings have not been systematically tested in environments where collaboration is not an important part of the professional culture. The relevance of LS to the specific challenges faced by Iranian English language teachers has not been fully established, while LS fosters enhanced pedagogical content knowledge and improves student learning outcomes (Takahashi & Yoshida, 2004).

Despite recent efforts to improve professional development initiatives aimed at empowering teachers to enhance student learning outcomes in Iran (Azari Noughabi & Amirian, 2021; Bonjakhi et al., 2024; Keyvanloo et al., 2023), significant gaps remain in the effectiveness and implementation of these programs. By emphasizing teacher agency and collaborative problem-solving, LS presents a promising model for sustainable professional development (Holmqvist, 2020). However, the present landscape of professional development in Iran often lacks the necessary structure and support to facilitate meaningful change. Ongoing evaluation and improvement of teaching practices are crucial since they are key factors in student learning success (Laoli et al., 2022). Inadequate teaching practices can severely hinder students' learning conditions, underscoring the urgent need for structured professional development approaches like LS.

Lesson Study (LS) has gained considerable attention globally over the past two decades, with various adaptations across Europe and the USA. However, its core principles remain deeply rooted in its origins in Japan, as well as in other parts of the Far East, such as Singapore, Hong Kong, and China (Lee, 2008). Significant challenges arise when attempting to implement LS in new educational contexts, while these global adaptations demonstrate the flexibility of LS. Although LS focuses on collaborative lesson planning, live lesson observations, and reflective dialogue, its success depends largely on the professional and cultural environment in which it is used.

One key problem with LS implementation outside of its cultural origin is the issue of teacher collaboration. In many Western and non-Asian educational settings, professional collaboration is not deeply embedded in the teaching culture. The UK version of LS, as outlined by Dudley (2013), shows

that teachers typically focus on a small number of case students and follow a cyclical process of research lessons, review, and planning meetings. However, even in the UK, the success of LS is not guaranteed without significant adaptation to the local teaching culture and practices. This raises concerns about how LS can be effectively implemented in educational systems that do not value collaborative professional development, such as in the Iranian context, where teachers often work in isolation.

While the fundamental goal of LS is to enhance teaching knowledge and improve student learning outcomes, literature largely focuses on its benefits rather than the barriers to its implementation. Takahashi & Yoshida (2004) emphasize the capacity of LS to promote reflective teaching practices and create opportunities for professional growth. However, these benefits are highly contingent on institutional support and the professional willingness of teachers to participate in LS cycles. Without addressing these contextual factors, LS may be viewed as an idealized model that is difficult to realize in less collaborative teaching environments.

Furthermore, while LS's adaptability has been widely praised, its implementation across different contexts often fails to account for critical variables such as resource availability, teacher workload, and the specific needs of foreign language teachers. As practiced in Dudley's (2013) UK model, the research lesson cycles might not fully translate into countries like Iran, where teaching resources are often limited and teachers face additional barriers to collaboration and professional development.

There has been another challenge showing that most studies on LS (e.g., Khalid et al., 2016; Schipper et al., 2017) tend to focus on its positive effects on teachers' knowledge and professional collaboration, which often neglect to critically examine its limitations in diverse educational settings. For example, while LS fosters dynamic lesson design and reflective practices, its ability to address the specific needs of foreign language instruction, especially in underdeveloped educational contexts, remains underexplored. Studies like Coşkun (2017) suggest that LS can support foreign language teachers in continuously improving their instructional practices; yet, little empirical research has been conducted on the logistical and cultural hurdles that must be overcome for LS to be effective in non-Western or non-Asian settings.

Thus, while LS offers a well-structured approach for enhancing instructional practices, its implementation needs careful consideration of cultural and contextual factors that may hinder its success in places like Iran. The current body of research, although rich in its depiction of LS's potential, fails to address these critical challenges, leaving a significant gap in the literature. This study seeks to fill this gap by investigating the feasibility of LS in Iranian English language institutes, where professional collaboration is not yet institutionalized. Understanding these barriers and proposing practical adaptations for LS in the Iranian context will contribute to a more nuanced understanding of how LS can be applied in diverse educational environments.

The literature on LS has increasingly recognized its potential as a collaborative professional development model for teachers, particularly in enhancing pedagogical practices and fostering a supportive teaching community. One significant study by Altınsoy (2023) in Turkey showed that LS effectively reduces isolation among novice teachers by promoting collaboration through cycles of lesson planning, observation, and reflection. In this study, participants exhibited significant growth in their pedagogical knowledge and collaboration skills, which suggests that LS can effectively bridge

the gap between theory and practice in teacher education, especially for pre-service teachers. However, this research mainly focused on pre-service educators, leaving unanswered questions regarding the impact of LS on experienced teachers who may have different professional development needs.

Building on these insights, Willems and Van den Bossche (2019) synthesized various studies on LS and confirmed its effectiveness in enhancing teacher knowledge, skills, and reflective practices. Their findings underscored the role of LS in creating collaborative environments that could be beneficial for both novice and experienced teachers. Nonetheless, they identified a critical gap in long-term research regarding the sustainability of LS practices across diverse educational settings, indicating a pressing need for empirical studies that explore the scalability and lasting effects of LS.

Vermunt et al. (2019) contributed to this discourse by examining the qualitative effects of LS on teacher learning and collaboration, focusing on reflective discussions during lesson planning and delivery. Their findings revealed that LS deepens teachers' understanding of student thinking and enhances instructional strategies, particularly through video analysis as a tool for refining teaching practices. However, they also pointed out the logistical challenges teachers face when employing digital tools for LS, especially in resource-constrained environments.

Further research has emphasized LS's potential to integrate with modern educational approaches such as STEM education. Aykan and Yildirim (2022) noted that LS improved STEM lesson planning and teaching quality during the COVID-19 pandemic. Their findings demonstrated that LS could enhance both pedagogy and content knowledge in distance learning contexts. However, this study did not thoroughly explore how LS can be sustained beyond crisis-driven distance learning or how its benefits can be scaled across various subjects and educational systems.

Laoli et al. (2022) examined challenges faced by novice English teachers and found that LS could alleviate issues like lack of confidence and reliance on traditional teaching methods. This research leaves open questions about how experienced educators might benefit from LS and how it could be adapted to environments lacking a collaborative culture, while emphasizing the importance of structured support systems for novice teachers' long-term success.

Although studies like Richit et al. (2021) and Durkin (2023) provide valuable insights into the collaborative nature of LS, there remains a noticeable gap regarding the adaptability of LS in non-collaborative, resource-constrained environments such as Iran. Most research tends to focus on the positive aspects of LS without critically addressing the logistical and cultural barriers impeding its implementation in such regions.

Research consistently underscores LS's role as a collaborative professional development model. For instance, Sabbagh et al. (2023) developed a model for managing LS in secondary schools that emphasizes mutual learning through qualitative research circles involving questioning, planning, action, observation, and reflection. This study lacks empirical evidence regarding the long-term sustainability of LS within non-Western educational systems where professional collaboration is less common, while it highlights the potential for peer support and practical experience exchange to enhance instructional quality.

In the EFL context, Zahedi et al. (2022) investigated LS's role in improving classroom education specifically within elementary schools in Iran. Their findings suggest that LS can improve teaching quality by fostering teacher collaboration; however, this research is limited to early educational settings and does not address whether similar benefits apply to higher education or foreign language instruction contexts like EFL teaching.

Building upon these insights, Farvardin et al. (2022) explored the impact of LS on both teachers' professional development and learners' grammatical knowledge. Their study demonstrated that LS enhances teacher collaboration and professional growth while supporting its adaptability to EFL settings. However, they did not completely address how LS can be implemented in systems that lack institutional support or where teachers face resource constraint issues prevalent in many non-Western educational systems.

Regardless of the empirical evidence supporting the benefits of LS, a significant gap exists in qualitative research focusing on institutes' managers, particularly through semi-structured interviews. Most existing studies have primarily focused on teachers' perceptions while largely neglecting the views of managers. This oversight is vital, as understanding managers' perspectives could provide deeper insights into the institutional support mechanisms necessary for the effective implementation of LS practices. Furthermore, while empirical research consistently stresses the advantages of LS, many studies overlook the challenges associated with its application in different cultural and institutional contexts. This is particularly relevant in regions like Iran, where a lack of a strong collaborative teaching culture may impede the adoption of LS. Addressing these gaps, this study aims to investigate the feasibility of implementing LS in Iranian English language institutes, focusing on identifying barriers and potential adaptations required for successful implementation.

There remains a significant gap in the literature concerning the perceptions of institute managers on LS, while teachers' perceptions have been extensively researched across various topics (e.g., Amirian & Azari Noughabi, 2017; Atai-Tabar et al., 2024; Richards et al., 2001), including their beliefs about LS (Khokhotva & Elexpuru Albizuri, 2020; Mayrhofer, 2019; Purwanti & Hatmanto, 2019) and how it shapes their professional learning opportunities. The current study aims to address this gap by exploring the perceptions of Iranian English Language Teaching (ELT) institute managers about the feasibility of applying LS in Iran to understand the cultural and contextual challenges that may influence its adoption. This is significant because, without clear evidence of its applicability from those in managerial positions, LS risks being disregarded as an effective professional development tool in Iran. The research questions that guided this study are as follows:

RQ1. What are Iranian English language institute managers' perceptions of the recognition of Lesson Study as a method for professional development in Iran?

RQ2. What are Iranian English language institute managers' perceptions of the effectiveness of Lesson Study as a method for professional development in Iran?

RQ3. What are Iranian English language institute managers' perceptions of the feasibility of Lesson Study as a method for professional development in Iran?

2. Methodology

This section includes the description of the participants, the instruments, and the data collection procedure.

2.1 Participants and Setting

This study employed a qualitative design, utilizing descriptive conceptual content analysis to identify concepts or ideas in the qualitative data and examine the study's objectives. The research was conducted in Mazandaran and Khorasan Razavi provinces. The participants consisted of six exclusively male institute managers aged 40–60, five Ph.D. holders, and one M.A. holder in ELT, selected through purposive sampling. This group included experienced ELT teachers with a minimum of 10 years of teaching experience and 5–30 years of management experience.

Purposive sampling is a kind of qualitative research technique that is used by researchers to deliberately choose participants who have specific characteristics relevant to the research focus. This ensures that the data collected is rich and relevant to the study's objectives. Additionally, it facilitates a deeper exploration of topics that may be overlooked in broader sampling methods. This leads to more detailed and contextually rich data (Creswell, 2016). Efforts were made to ensure diversity by including managers from institutes of varying sizes and operational models to provide a range of perspectives. Although the sample size was small, it was deemed appropriate for an exploratory feasibility study aimed at generating deep insights into the potential implementation of LS in Iranian context.

2.2 Instrumentation

Institute managers were chosen in this study for semi-structured interviews because of their experience in both management and teaching, as well as their understanding of institutional challenges related to professional development. These interviews were the primary data collection instrument conducted with the selected institute managers and lasted between 30 to 60 minutes. Among these six interviews, two took place face-to-face, while the remaining four were conducted online due to COVID-19 restrictions. In-depth exploration of topics is among the benefits of semi-structured interviews, which keep a consistent framework and allow for both reliable data collection and rich qualitative insights (Creswell, 2016). The mentioned interviews consisted of 13 open-ended questions designed to explore managers' views on cognition, perception, and the feasibility of LS implementation in their institutes. These questions were developed based on a review of the relevant literature and validated by seven experts in ELT to ensure their validity. Example questions included:

"Have you studied or researched Lesson Study? If so, could you share your insights?"

"What is the fundamental difference between Lesson Study and other methods of improving education?"

"Do financial issues hinder the implementation of Lesson Study?"

"Have you ever conducted or do you plan to conduct Lesson Study at your institute?"

2.3 Research Procedure

The study was conducted over a period of 5 to 6 months and was designed as a feasibility study without an experimental or control group. Participant selection involved approaching 30 institutes across various cities. Out of these, 16 institutes refused to participate without considering the topic. Ultimately, 14 institutes consented to cooperate; however, eight of these withdrew after being informed about the specifics of the study, showing a lack of recognition of LS as an important professional development method. This left six English language institutes that agreed to participate in the study. The remaining six institute managers participated in the semi-structured interviews to provide insights into their cognition, perceptions, and challenges related to LS. The interview protocol of this study, adopted from Jacob and Furgerson (2012), begins with an introduction that briefly explains the purpose of the study and obtains informed consent from institute managers as the participants. Following this, warm-up questions are posed to gather general information about the participants' backgrounds and experiences related to education and professional development. The core questions focus on open-ended inquiries regarding previous knowledge or research on LS, perceptions of its effectiveness as a professional development method, challenges faced in implementing LS, and plans for future implementation. Exploring questions is encouraged to elicit elaboration on responses, which allows for deeper insights based on the participants' answers. The interviews conclude with a summary of key points discussed, thanking the participants for their time and contributions. Additionally, the interviews were recorded and subsequently transcribed for analysis. The transcribed interviews underwent descriptive conceptual content analysis using MAXQDA software, which facilitated the identification of recurring themes and codes.

Through deductive descriptive coding, 14 codes were extracted from the data, including belief, cognition, efficiency, implementation, preference, obstacles, financial issues, voluntary implementation, permission, class observation, interaction, collaboration, teacher, and teaching. These codes were derived from a thorough review of the transcribed interview data and aligned with the study's research questions as well as relevant theoretical frameworks and empirical studies on LS. The codes were analyzed to determine their significance in relation to the research objectives. The frequency of each code's occurrence was quantified using SPSS version 26 to enhance the rigor of the content analysis. This quantitative component provided additional insights into the prominence of specific themes within the dataset. By integrating both qualitative insights and quantitative frequency counts, the study ensured a comprehensive exploration of the feasibility of LS implementation in Iranian English language institutes.

3. Data Analysis

To address the first research question concerning the perceptions of Iranian Language Institute managers toward recognizing LS as a method for professional development, interviews with managers of these institutes and analysis of the collected data revealed several key codes. Among these codes, cognition, efficiency, and implementation emerged as the most critical in addressing the research questions. Notably, the code "cognition" was identified as particularly significant in understanding the managers' perceptions of LS. The insights shared by interviewees regarding their understanding of LS are summarized in Table 1, which illustrates their varying levels of awareness and recognition of this professional development approach.

Table 1*The ideas given by the interviewees on their cognition of LS*

	Cognition
Interviewee One	<ul style="list-style-type: none"> -Yes, I know LS. -I have a personal interest in this topic. -I have studied and conducted a workshop on the subject.
Interviewee Two	<ul style="list-style-type: none"> -I don't know it academically. -We do not have academic knowledge, but I have some general information from my university days. -LS is carried out based on the needs of the teachers and the class, and there is no fixed method.
Interviewee Three	<ul style="list-style-type: none"> -It is an effective method, and based on the explanations given, it is very good. -The main difference is that it is natural and derived from the practice to theory. -I have also experimented with it, using several professors in teaching. This creates a valuable reflection among teachers in the class. -I may not know it academically, but according to the principle of LS (teachers' cooperation to solve class problems), this topic was my idea before establishing the institute.
Interviewee Four	<ul style="list-style-type: none"> -I have personally studied peer observation and mentoring, and I have read some materials, but I have no knowledge and no study about LS, and I have not seen it at any institutes, even in famous institutions like the ILI. -In my opinion, the difference between Lesson Study and other methods is that the teachers can choose the right method based on their own problem with the cooperation of other teachers and the needs of the class. -I did not do it in my institute and my answer is negative. But in my opinion, based on the explanations given, it can be very effective.
Interviewee Five	<ul style="list-style-type: none"> -Yes, there is a difference. LS is a type of participatory approach that involves participation among teachers. -I also wrote a thesis on LS, and it is like action research, so the teacher must have the ability to reflect with others. -Teachers observe each other and use their experiences to solve classroom problems with the aim of improvement. As for me, I have a very positive opinion about it.
Interviewee Six	<ul style="list-style-type: none"> -I am familiar with LS, but it has not been implemented in a structured manner in the class until now. However, I believe that having someone come and observe the class and provide feedback would be beneficial. -There is something similar to LS called action research, which addresses the issues present in the class. It is worth noting that action research emphasizes solving one's own problems. -The main distinction between LS and action research is that LS is conducted in groups, where observations are made, and the findings are used to make adjustments to the lessons. On the other hand, action research can be carried out individually, for personal and classroom improvement purposes. -In my opinion, LS falls under the category of CPD (Continuous Professional Development). Attending CPD activities is a way to enhance one's career.

Based on the data received from interviews, the majority of the institute managers showed a limited or incomplete understanding of LS. Out of the six institute managers interviewed, four (66.7%)

showed a distorted familiarity with LS, indicating that while they had heard of the concept, their knowledge was inaccurate or superficial. Two managers (33.3%) had either no or minimal knowledge of LS, reflecting a significant gap in awareness. Only one institute manager (16.7%) had a thorough understanding of LS, having studied and applied it in a comprehensive manner. This pattern highlights a general lack of cognition about LS among Iranian English language institute managers. The use of LS in Iran faces significant challenges due to limited awareness and distorted familiarity among those who would use it, even though LS has proven to be an effective professional development approach in other contexts. These findings highlight the need to educate managers on LS, which would be an important first step toward fostering its use in the Iranian educational system. To address the second research question concerning institute managers' perceptions of the effectiveness of LS for their professional development, the researcher focused on analyzing data related to the code "efficiency." This analysis involved transcribing the interviewees' insights regarding their perceptions of the efficiency of LS, which are specified in Table 2. The results show how managers evaluate the impact of LS on their professional growth and highlight their beliefs about its effectiveness as a developmental approach within their institutes.

Table 2

The ideas given by the interviewees on the efficiency of LS

	Efficiency
Interviewee One	<ul style="list-style-type: none"> -I strongly believe in it. -It is definitely very effective. -And it is believed to be efficient. -It will solve many problems in the class and promote the teacher and education.
Interviewee Two	<ul style="list-style-type: none"> Yes. If it can be fully implemented, it will be an effective method.
Interviewee Three	<ul style="list-style-type: none"> -LS creates a very positive impression among the teachers in the class. -And, based on the explanations given, it is very beneficial. -And if I were to implement it and the problems of the institutes were reduced, I would bring the entire classes and teachers to it.
Interviewee Four	<ul style="list-style-type: none"> -I firmly believe that the concept of Lesson Study (LS) needs to be cultivated within our educational culture -LS is inherently subjective, but based on the explanations provided, I have strong faith in its value. -Teachers collaborating and supporting one another is crucial, and I am fully convinced that LS can be 100% effective. -Furthermore, if implemented correctly, I believe LS has the potential to address many educational challenges.
Interviewee Five	<ul style="list-style-type: none"> -The goal is to improve and address problems in LS. -Professors can observe each other's classes in groups and collaboratively solve problems, which -I believe can be very beneficial. -LS can be highly advantageous. Teachers can address these problems by working together.
Interviewee Six	<ul style="list-style-type: none"> -LS is very practical; it can be used to solve problems in class. -I think it is very effective, but if it is done correctly, it is wonderful. If you remove the hassle from it, the results are very good and, at the same time, the atmosphere of the classroom and the institute will be very good.

The data revealed that five out of six institute managers (83.3%) viewed LS as a useful and efficient method for teaching skills development. These managers expressed a strong belief in the effectiveness of LS to foster professional development, reflection, and collaboration. They emphasized that, if used appropriately, LS could significantly enhance teacher development and classroom instruction in their institutes. However, one manager (16.7%) did not consider LS relevant or effective for their context. This could be due to either a lack of understanding of LS or challenges in adapting it to the specific needs of their institute. Overall, the results reveal that even though most of the managers believe in the efficiency of LS, its successful implementation would rely on addressing gaps in understanding and ensuring that all managers and teachers are fully informed of its benefits. Without sufficient knowledge and awareness, LS may not be applied effectively in the Iranian English language institutes' professional development frameworks. Therefore, raising awareness and providing training on LS is needed to enable its implementation and maximize its potential for teacher development. In addressing the third research question involving the feasibility of implementing LS as a method of professional development in Iranian English language institutes, the relevant code identified was "implementation." This code contains various insights about the practical application of LS. The ideas expressed by the interviewees regarding their experiences and strategies for implementing LS are transcribed and analyzed in Table 3.

Table 3

The ideas given by the interviewees on their Implementation of LS

Implementation	
Interviewee One	<p>-Our teaching method at the institute is for newcomers to observe at least 10 classes and write a report on them after completing the TTC course.</p> <p>This approach has been very helpful, as one of the new teachers who had many problems in the classroom has made significant progress.</p> <p>-Our teachers observe each other's classes to address any questions they may have, such as how the teacher is managing the class and what strategies they are using. These questions and answers have helped solve many problems.</p> <p>-For their own benefit, teachers can invite other teachers to observe their classes and provide feedback and comments. Alternatively, a less experienced teacher can attend an experienced teacher's class to learn new things. Teachers who teach the same class can also observe each other and brainstorm together.</p>
Interviewee Two	<p>-In our teacher recruitment system, after the new teachers are trained, they must attend several classes and observe other teachers to learn the institute's processes.</p> <p>-We had a course where the training was intensive and the class content was extensive, so the teachers communicated with each other and discussed the issues during these meetings.</p> <p>-In our institute, there are regular meetings where teachers are asked for their opinions on various aspects of teaching, such as departmental matters, class management, and scheduling.</p> <p>-Whether it was regarding educational or management issues, there were instances where experienced teachers provided assistance to those who were less experienced and offered solutions on how to handle various classroom situations.</p>

Interviewee Three	<p>-We didn't have LS, but one issue was very important to me: good support was formed among the language learners, which was very satisfying.</p> <p>-Naturally, yes, I did some parts of LS with the discussions and observations that the professors made together, not only in the field of educational processes, but also in other issues such as class and time management, they were able to reduce their problems a lot.</p>
Interviewee Four	<p>-No, I have not done it yet, but if the institute undergoes significant development and the number of students increases, I will definitely do it.</p> <p>-I have not done it yet, but I can confidently say that if LS is implemented in my institute, it will undoubtedly improve the educational process.</p>
Interviewee Five	<p>-Some parts of LS have been used.</p> <p>-Many problems in the classes have been solved thanks to the experiences of experts, which have been very helpful. However, as I mentioned, the LS study has not been fully implemented.</p>
Interviewee Six	<p>-We had some parts of the Lesson Study, such as peer observation and mentoring, and we involved the teachers in the discussions. However, it has not been applied in a structured way in the class until now. However, it is possible for someone to come and observe the class and give feedback as part of the course research.</p> <p>-Because it is not conventional, there may be more problems than benefits. However, if it is done correctly, it can be wonderful.</p> <p>-We have had some collaborations, but we have not had an academic and structured LS.</p> <p>If we want to implement it in a structured way, we can choose several teachers to participate. It must be culturalized.</p>

While LS can present a lot of advantages to students and teachers by promoting a culture of continuous improvement and professional growth, the data from the interviews illustrate a significant gap in its applicability in Iranian English language institutes. Out of the six institute managers interviewed, only one manager (16.7%) with prior knowledge and implementation experience expressed a positive view regarding the feasibility of implementing LS in the institute. The remaining five managers (83.3%), not knowing LS prior to the interview, considered it irrelevant and incompatible with their context due to fundamental knowledge gaps about the LS concept and institutional barriers such as financial obstacles, restrictive policies against observation, and the absence of collaborative frameworks that must be culturalized. This perception of LS as "not relevant" among initially unaware managers means that before the successful adoption of LS in Iranian institutes, much has to be done to raise awareness and provide training for institute managers and teachers. Without a foundational understanding of LS and its advantages, the likelihood of successful implementation remains low. Therefore, while LS offers considerable potential, it is not currently viewed as applicable within the current structure of Iranian English language institutes without systemic changes to address knowledge deficits and institutional barriers.

4. Discussion and Conclusion

The aim of the current study was to explore the feasibility of implementing LS as a professional development tool for English as a Foreign Language (EFL) teachers in Iranian English language

institutes. It sought to determine whether LS is recognized by institute managers as an effective method for teacher development and whether it is suitable to enhance EFL teachers' professional development in Iran.

The findings indicate a relationship between an understanding of LS and its application for teachers' professional development. Managers who were aware of LS recognized its potential to foster reflective practices, enhance collaborative lesson planning, and improve classroom teaching. These results are consistent with the work of Yoshida (1999), Fujii (2019), and Arslan (2019), where LS is used to enhance teaching practices through peer collaboration and constructive feedback. However, LS is still relatively unfamiliar to many educators outside of Japan and in some regions like Iran, where limited knowledge of its implementation strategies poses challenges. This observation aligns with the findings of Bayram & Bıkmaz (2021), who believe that increased adoption of LS would be achievable if teachers and administrators were professionally trained and supported for effective implementation.

In the Iranian context, the study highlighted that even though LS has the potential to significantly influence the professional development of EFL teachers, its adoption is hindered by a lack of teacher awareness and institutional support. The findings reflect the work of Ashraf & Kafi (2017), Arslan (2019), and Derakhshan et al. (2020), who also identified LS as a powerful tool for teacher development, especially when educators fully understand its processes. Managers who had some insight into LS believed it could improve teacher collaboration and instructional quality, provided that teachers were comprehensively informed about its benefits. However, the absence of a collaborative professional culture remains a significant obstacle. In Iran, EFL teachers are predominantly isolated and are not provided with sufficient opportunities to engage in reflective and collaborative practice, which is central to the success of LS.

Despite the evidence of LS's benefits in enhancing pedagogical content knowledge and improving classroom practices (Cerbin & Kopp, 2006), its practical application in Iran is not well developed. Studies by Alipour (2018) and Haghighifard & Marzban (2016) confirm that LS is not utilized to its maximum potential in Iranian language institutes due to a lack of institutional support, limited resources, and the absence of a collaborative professional development culture. While some managers recognize the potential of LS, others view it as irrelevant or challenging to implement, citing these constraints as significant barriers. This finding is consistent with research by Richit et al. (2021) and Durkin (2023), emphasizing that leadership plays a pivotal role in either facilitating or hindering the adoption of new educational practices like LS.

To overcome such challenges, there is a need to raise awareness about the benefits of LS and to professionally train managers and teachers. Establishing institutional structures that encourage collaboration could significantly enhance the effectiveness of LS in Iranian English language institutes. Willems and Van den Bossche (2019) indicate the need for more extensive, long-term research to assess the sustained impact of LS in various educational settings, suggesting that without suitable leadership and institutional support, the benefits of LS may remain limited.

Additionally, the results of this study indicate that the applicability of LS in Iran is constrained not by the inherent limitations of the concept, but by the reluctance of managers and teachers to embrace it fully. This is consistent with the findings of Alipour (2018) and Haghighifard & Marzban (2016), who also identified a gap between the potential benefits of LS and its actual implementation in Iranian English language institutes. This reluctance highlights the need for cultural and contextual adaptation of LS to better fit the Iranian educational system.

This research highlights that LS is a valuable professional development initiative through which teachers collaborate to plan, observe, and reflect on lessons. LS provides a means for teachers to carry out classroom research, helping them identify areas for improvement and develop more effective teaching strategies. Its inherent flexibility and emphasis on collaborative inquiry make it an effective method for enhancing teaching practices and learners' outcomes in various educational contexts, including Iranian English language institutes.

To promote the use of LS in these institutes, it would be ideal to share success stories of English language teachers who have effectively implemented LS in their classrooms. Organizing training sessions on LS can equip teachers with the necessary knowledge and skills for its effective implementation. These sessions should incorporate hands-on activities, collaborative lesson planning, and reflective activities to deepen participants' understanding of LS principles.

Additionally, establishing partnerships with universities and other educational institutions can facilitate the promotion of LS as a professional development method, which may eventually lead to joint training programs or the integration of LS into existing professional development initiatives. Raising awareness among teachers and institutional managers about the benefits of LS is extremely crucial for its effective implementation. A comprehensive understanding of LS can foster strong support among educators, while a lack of recognition remains a primary barrier to its use.

By providing targeted information and training, institutes can strengthen beliefs in the efficacy of LS and establish a good foundation for overcoming difficulties. Furthermore, LS practitioners and institute managers implementing LS in diverse teams need to be aware of practical factors such as team composition, students' conversational skills, and the learning characteristics of participating schools. Addressing these considerations will help ensure that LS fosters successful collaboration and achieves its intended outcomes, ultimately contributing to the professional growth of EFL teachers and enhancing student learning experiences.

The findings of this study carry considerable implications for professional development in the Iranian educational setting. LS is a promising approach to enhancing teaching practices and fostering collaboration among educators. For successful implementation, it is essential that institutes provide adequate resources, time, and support for collaborative practices like LS, with the active support of management. Comprehensive training programs must be developed to equip teachers with the necessary skills and knowledge to engage effectively in LS, along with practical activities such as lesson planning, observation, and reflective exercises. Additionally, LS must be culturally and contextually adapted to align with the characteristics of Iranian English language institutes, addressing potential resistance to collaboration and fostering environments beneficial to open discussions. Access to qualified facilitators who understand LS principles is essential for guiding collaborative processes, and institute managers need to be trained so that they can support LS initiatives within their institutes. Finally, the integration of LS into broader institutional policies and professional development programs will help ensure its sustainability and long-term impact on teacher development.

This study faced several limitations that need to be mentioned. Firstly, the sample size was limited to six institute managers, which restricts the generalizability of the findings; a larger and more diverse sample in future research could provide a more comprehensive view of LS implementation. Additionally, while the focus on managers' perspectives was in line with the research objectives, it excluded teachers' direct experience of LS, thereby limiting what is known about its practice in the classroom. The geographic scope of the study was also limited to two provinces, which may not

provide an encompassing view of the Iranian context; extending the research coverage to additional areas would enhance its applicability. Finally, logistical constraints due to the COVID-19 pandemic necessitated online interviews for four participants, which may have impacted the depth of data collection and the richness of responses.

Given these results, there are some directions for future research to explore and expand the potential of LS as a professional development model. First, more empirical studies are needed to examine the long-term effects of LS on teachers' performance and students' outcomes in Iran. As Vermunt et al. (2019) note, reflective discussions of lesson planning and teaching can deepen teachers' understanding of student learning; however, logistical challenges such as resource constraints and the availability of digital tools must be addressed if LS is to be fully effective. Additionally, action research could be conducted to pilot LS programs in selected Iranian institutes and form a foundation for extensive implementation and help to refine LS practices for under-resourced environments. For instance, Altinsoy (2023) demonstrated that LS has the potential to bridge the gap between theory and practice in pre-service teacher education, but further research should examine how LS can be adapted for in-service teacher development, particularly in resource-constrained settings like Iran. Future studies should also explore the impact of LS on students' learning outcomes in English language classrooms through quantitative or mixed-method approaches.

Furthermore, it would be beneficial to examine teachers' experiences and perceptions towards LS to provide a more thorough understanding of its challenges and benefits. Research on how LS can be culturally adapted to address specific challenges in Iranian educational contexts, such as fostering collaboration in traditionally isolated teaching environments, would also be valuable. Lastly, conducting longitudinal studies to measure the effects of LS on teachers' motivation, retention, and job satisfaction could provide deeper information on its long-term effects.

Acknowledgements

The researchers would like to express their heartfelt gratitude to those participants who willingly and voluntarily attended this study.

References

- Alipour, S. (2018). The effect of lesson study on Iranian EFL teachers' perception of teaching. *Language Teaching Research Quarterly*, 5, 45-58. <https://doi.org/10.32038/ltrq.2018.05.03>
- Altinsoy, E. (2023). Lesson Study: Does it work for personal development of pre-service ELT teachers? *Gist: Education and Learning Research Journal*, 26, 57-79. <https://doi.org/10.26817/16925777.1569>
- Amirian, M. R., & Azari Noughabi, M. (2017). An investigation into Iranian EFL teachers' perception of learner autonomy. *Journal of Language Teaching and Research*, 8(4), 807-816. <https://doi.org/10.17507/jltr.0804.22>
- Arslan, F. Y. (2019). The role of lesson study in teaching, learning, and professional development of EFL teachers in Turkey: A case study. *TESOL Journal*, 10(2), 1-13. <https://doi.org/10.1002/tesj.409>

- Ashraf, H., & Kafi, Z. H. (2017). Philosophy of education subscales: Predictors of EFL teachers' professional development. *Journal on English Language Teaching*, 7(1), 11-19. <https://doi.org/10.26634/jelt.7.1.11403>
- Atai-Tabar, M., Zareian, G., Amirian, S. M. R., & Adel, S. M. R. (2024). Relationships between EFL teachers' perceptions of consequential validity of formative assessment and data-driven decision-making self-efficacy and anxiety. *Journal of Applied Research in Higher Education*, 16(3), 919-933. <https://doi.org/10.1108/jarhe-04-2023-0169>
- Aykan, A., & Yildirim, B. (2022). The integration of a lesson study model into distance STEM education during the COVID-19 pandemic: Teachers' views and practice. *Technology, Knowledge and Learning*, 27(2), 609-637. <https://doi.org/10.1007/s10758-021-09564-9>
- Azari Noughabi, M., & Amirian, S. M. R. (2021). Assessing the contribution of autonomy and self-efficacy to EFL teachers' self-regulation. *English Teaching & Learning*, 45(1), 71-88. <https://doi.org/10.1007/s42321-020-00060-4>
- Bayram, İ., & Bıkmaz, F. (2021). Implications of lesson study for tertiary-level EFL teachers' professional development: A case study from Turkey. *Sage Open*, 11(2), <https://doi.org/10.1177/21582440211023771>
- Bonjakhi, M., Amirian, S. M. R., Davoudi, M., & Adel, S. M. R. (2024). The contribution of psychological well-being, L2 teacher grit, grit-s, and perceived organizational support to teacher continuous professional development: The mediating role of work motivation. *Language Related Research*. <https://lrr.modares.ac.ir/article-14-73962-fa.html>
- Cerbin W., & Kopp B. (2006). Lesson study as a model for building pedagogical knowledge and improving teaching. *International Journal of Teaching and Learning in Higher Education*, 18(3), 250-257.
- Coenders, F., & Verhoef, N. (2019). Lesson study: Professional development (PD) for beginning and experienced teachers. *Professional Development in Education*, 45(2), 217-230. <https://doi.org/10.1080/19415257.2018.1430050>
- Coşkun, A. (2017). The application of lesson study in teaching English as a foreign language. *İnönü Üniversitesi Eğitim Fakültesi Dergisi*, 18(1), 151-162. <https://doi.org/10.17679/inuefd.297845>
- Creswell, J. (2016). *Research design: Qualitative, quantitative, mixed-methods approaches* (5th ed.). Sage Publications.
- 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. <https://files.eric.ed.gov/fulltext/EJ1275848.pdf>
- Dudley, P. (2013). Teacher learning in lesson study: What interaction-level discourse analysis revealed about how teachers utilized imagination, tacit knowledge of teaching, and fresh evidence of pupils learning, to develop practice knowledge and so enhance their pupils' learning. *Teaching and Teacher Education*, 34, 107-121. <https://doi.org/10.1016/j.tate.2013.04.006>
- Durkin, M. S. (2023). The importance of lesson study. *Mathematics Teacher: Learning and Teaching PK-12*, 116(5). <https://doi.org/10.5951/MTLT.2023.0085>

- Farvardin, M. T., Abbasi Behbahani, L., & Ebrahimi Askari, F. (2022). Effect of lesson study on EFL learners' grammatical knowledge and teachers' perception. *Journal of New Advances in English Language Teaching and Applied Linguistics*, 4(1), 799-815. https://www.jeltal.ir/article_147011.html
- Fujii, T. (2019). Designing and adapting tasks in lesson planning: A critical process of lesson study. *Theory and practice of lesson study in mathematics: An international perspective* (pp. 681-704). Springer.
- Godfrey, D., Seleznyov, S., Anders, J., Wollaston, N., & Barrera-Pedemonte, F. (2019). A developmental evaluation approach to Lesson Study: Exploring the impact of lesson study in London schools. *Professional Development in Education*, 45(2), 325-340. <https://doi.org/10.1080/19415257.2018.1474488>
- Haghighifard, M., & Marzban, A. (2016). Teachers' viewpoints on the practical implementation of lesson study in EFL classrooms. *International Journal of Humanities and Cultural Studies (IJHCS)*, 3(2), 700-713.
- Holmqvist, M. (2020). Lesson study as a vehicle for improving SEND teachers' teaching skills. *International Journal for Lesson & Learning Studies*, 9(3), 193-202. <https://doi.org/10.1108/ijlls-05-2020-0022>
- Jacob, S. A., & Ferguson, S. P. (2012). Writing interview protocols and conducting interviews: Tips for students new to the field of qualitative research. *The Qualitative Report*, 17(42), 1-10. <https://doi.org/10.46743/2160-3715/2012.1718>
- Keyvanloo, M., Amirian, M. R., Vosoughi, M., & Bagheri Nevisi, R. (2023). Exploring the relationship among EFL teachers' critical thinking, autonomy and experience in public and private schools. *Journal of Modern Research in English Language Studies*, 10(1), 101-131. <https://doi.org/10.30479/jmrels.2022.17438.2085>
- Khalid, M., Abdullah, N. A. H., & Kamoludeen, A. (2016). Teachers' beliefs on the benefit of collaboration in lesson study. *IJUM Journal of Educational Studies*, 4(2), 1-20. <https://doi.org/10.31436/ijes.v4i2.115>
- Khokhotva, O., & Elexpuru Albizuri, I. (2020). Teachers' educational beliefs change through lesson study: Implications for school culture. *International Journal for Lesson & Learning Studies*, 9(4), 317-331. <https://doi.org/10.1108/IJLLS-04-2020-0018>.
- Laoli, A., Dakhi, O., & Zagoto, M. M. (2022). The application of lesson study in improving the quality of English teaching. *Edukatif: Jurnal Ilmu Pendidikan*, 4(2), 2238-2246. <https://doi.org/10.31004/edukatif.v4i2.2434>
- Lee, J. F. (2008). A Hong Kong case of lesson study—Benefits and concerns. *Teaching and Teacher Education*, 24(5), 1115-1124. <https://doi.org/10.1016/j.tate.2007.10.007>
- Lewis, C. C., Takahashi, A., Friedkin, S., Liebert, S., & Houseman, N. (2022). Sustained, effective school-wide lesson study: How do we get there? *Vietnam Journal of Education*, 45, 45-57. <https://doi.org/10.52296/vje.2022.178>

- Mayrhofer, E. (2019). Lesson study and teachers' beliefs: How a Bourdieuan perspective could make a difference. *International Journal for Lesson and Learning Studies*, 8(1), 19-33. <https://doi.org/10.1108/ijlls-11-2018-0091>
- Purwanti, E., & Hatmanto, E. D. (2019). Understanding EFL teachers' beliefs about lesson study and their knowledge development viewed from social cultural theory of Vygotsky. *English Language Teaching Educational Journal*, 2(2), 50-61. <https://files.eric.ed.gov/fulltext/EJ1283027.pdf>
- Richards, J. C., Gallo, P. B., & Renandya, W. A. (2001). Exploring teachers' beliefs and the processes of change. *PAC journal*, 1(1), 41-58.
- Richit, A., Ponte, J. P., & Tomasi, A. P. (2021). Aspects of professional collaboration in a lesson study. *International Electronic Journal of Mathematics Education*, 16(2), em0637. <https://doi.org/10.29333/iejme/10904>
- Sabbagh, F., Alizadeh, S., Jahed, H., sorani, R., & Bayanati, M. (2023). Designing a model for lesson study management in secondary schools. *Educational Leadership & administration*, 17(1), 83-106. <https://sanad.iau.ir/en/Article/857002>
- Schipper, T., Goei, S.L., Vries, S. & Veen, K. (2017). Professional growth in adaptive teaching competence as a result of lesson study. *Teaching and Teacher Education*, 68, 289-303. <https://doi.org/10.1016/j.tate.2017.09.015>
- Takahashi, A., & Yoshida, M. (2004). Ideas for establishing lesson study communities. *Teaching Children Mathematics*, 10(9), 436-443. <https://doi.org/10.5951/tcm.10.9.0436>
- Vermunt, J., Vrikki, M., Van Halem, N., Warwick, P., & Mercer, N. (2019). The impact of lesson study professional development on the quality of teacher learning. *International Journal of Educational Research*, 81, 61-73. <https://doi.org/10.1016/j.tate.2019.02.009>
- Willems, I., & Van den Bossche, P. (2019). Lesson Study effectiveness for teachers' professional learning: A best evidence synthesis. *International Journal for Lesson and Learning Studies*, 8(4), 257-271. <https://doi.org/10.1108/IJLLS-04-2024-0031>
- Yoshida, M. (1999). *Lesson study: An ethnographic investigation of school-based teacher development in Japan* (Unpublished doctoral dissertation). University of Chicago.
- Zahedi, S., Shafizadeh, H., Rezghi Shirsavar, H. (2022). Identifying the dimensions and components of promoting the lesson study in the elementary school. *Islamic Life Style*, 5(4), 435-444. https://www.islamiilife.com/article_186234_086236b7945559ebd320b63ca9daab30.pdf

Appendix

Interview Questions

1. Have you ever studied or researched Lesson Study? If so, could you share your insights?
2. What is the fundamental difference between Lesson Study and other methods of improving education?
3. Do you believe Lesson Study is an effective method for professional development?
4. Have you ever conducted or do you plan to conduct Lesson Study at your institute?
5. What obstacles do you think hinder the implementation of Lesson Study in your institute?
6. Do financial issues hinder the implementation of Lesson Study?
7. Is it possible to carry out parts of Lesson Study work spontaneously and voluntarily? (In cases where financial issues are considered to be an obstacle to implementation)
8. Are teachers willing to participate in a method of professional development that does not offer awards, points, letters of appreciation, or similar incentives? With what motivation or purpose?
9. If Lesson Study has been implemented in your institute, how effective was it in solving problems and improving the educational process?
10. If Lesson Study is implemented in your institute, to what extent has it solved the issues raised in the classroom and education? Please give an example if possible.
11. Have your teachers ever been allowed to observe each other's classes, or do they have permission?
12. Do the teachers at your institute sit together and discuss financial, educational, and disciplinary issues?
13. For the implementation of Lesson Study at your institute, is your emphasis on improving teaching (education) or on teachers?

We cordially encourage you to submit your work to our **Journal of Second Language Pedagogy**, which actively promotes the growth of scholarship in the field of second language education. Our journal is dedicated to the dissemination of cutting-edge research and the cultivation of critical discourse among scholars and practitioners. By contributing to this platform, authors not only expand the existing body of knowledge in the field, but also inspire further research and innovation in second language pedagogy. Through conducting thorough empirical research, perceptive theoretical investigations, and useful applications, you add to the body of knowledge that improves and informs language instruction.

JSLP Volume1 , Issue 3
2025