




<https://doi.org/10.22077/ali.2023.6251.1012>

# Productive Vocabulary Size Test as a Predictor of L2 Learners' Success in Academic Writing Skill

Atika Etemadzadeh<sup>1\*</sup>  
Khairi Izwan Abdullah<sup>2</sup>

<sup>1</sup>Ph.D. Candidate in Applied Linguistics, Universiti Teknologi Malaysia (UTM), Malaysia

<sup>2</sup>Associate Professor in Applied Linguistics, Universiti Teknologi Malaysia (UTM), Malaysia

## ARTICLE HISTORY

Received: 14 April 2023  
Revised: 14 May 2023  
Accepted: 21 June 2023  
Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [aticca.etemad@gmail.com](mailto:aticca.etemad@gmail.com)

## ABSTRACT

In this study, an attempt has been made to investigate if vocabulary tests may predict L2 learners' writing scores in IELTS. 131 postgraduate international students studying at Universiti Teknologi Malaysia (UTM) who obtained the minimum score of 6 in IELTS as a requirement for enrollment were participants of the study. Two vocabulary size tests (PVL & Lex30) together with a test of vocabulary depth (WAT) were given to participants. Then, participants' scores from three tests were correlated with their scores in the writing section of IELTS. The results of the study revealed a significant correlation between productive vocabulary size and academic writing skills. Moreover, the findings of Multiple Regression indicated that the productive vocabulary levels test (PVL) could be a good predictor for IELTS academic writing. Furthermore, scatter plots revealed a threshold of 2000 words for an L2 learner to get 7 in IELTS academic writing. The findings drawn from the study enable teachers to predict their learners' competency in writing skills through various vocabulary tests.

**KEYWORDS:** Productive vocabulary; Vocabulary size; Academic writing; IELTS

## 1. Introduction

Investigations into the vocabulary size of second language (L2) learners have greatly increased our understanding of the importance of vocabulary in language use. One consistent finding is that vocabulary size – the number of words a person knows – is a good predictor of proficiency in a second language. It has been found that vocabulary size in English correlates highly with general language proficiency (Milton, 2009; Nasir et al., 2017), reading comprehension (Aliabadi et al., 2022; Qian, 2002), writing ability (Alsager & Milton, 2016; Laufer & Nation, 1995), examination grades on listening, reading and writing papers (Afshari & Tavakoli, 2017; Staehr, 2008) and academic success (Morris & Cobb, 2004). The relationship between vocabulary size and language performance is so close that Alderson (2007), after an analysis of the correlation between a test battery of vocabulary skills and language proficiency, concluded that "...the size of one's vocabulary is relevant to one's performance on any language test, in other words, ...language ability is to quite a large extent a function of vocabulary size" (p. 88). Similar findings on word frequency and vocabulary size have also been reported for French as a second language (Milton, 2009).

If vocabulary size is closely related to second language ability, learners' vocabulary size could be used to predict how well a learner may perform in second language examinations. We believe this predictive potential of vocabulary size is worth exploring for two reasons. Firstly, learner performances in second language examinations (for example, IELTS, TOEFL, etc.)

are crucial for many of the stakeholders – learners, teachers, parents, and school administrators. The results are used for important decision-making related to acceptance into programs of study abroad, award of sponsorship, or simply, learner readiness for study in a second language. If vocabulary size is found to be a good predictor of performance in second language examinations, tangible vocabulary size targets could be set for both learners and teachers.

### 1.1. Challenges and contextual constraints initiating the study

As Nation and Laufer (2001) indicate "vocabulary is not usually learned for its own sake" (p. 2), an important goal of learning vocabulary is to use vocabulary knowledge communicatively. On the importance of the role of vocabulary in communication, Wilkins (1972) notes that "Without grammar very little can be conveyed, without vocabulary nothing can be conveyed" (p. 111). All people involved in language teaching/learning processes agree on the essential role of vocabulary as a part of language learning mastery. This role is more apparent when learners try to use their language in real contexts and it is usually evident in "learners carry around dictionaries and not grammar books" (Schmitt, 2010, p. 4). One crucial aspect of communication is writing in an academic context. Foreign students, studying in international universities, particularly for postgraduate degrees, in which students need to write academic papers in their related areas, must possess and use many academic skills. Certain parts of these academic endeavors include writing assignments, dissertations, and/or academic papers. To be successfully engaged in these activities, foreign students need an excellent command of vocabulary knowledge and academic writing skills. In this matter, Mohan and Lo (1985) highlight the importance of vocabulary in academic writing and believe that some of the difficulties that L2 learners usually face in academic writing are problems at the sentence level, some of which are grammatical but most are lexical. Ultimately, it has been found that non-native students who are equipped with richer word knowledge tend to have better performance in academic education.

Further, from second language learners' points of view, L2 learners have continually reported their lack of vocabulary knowledge as the fundamental aspect of their writing quality (Lee & Rethinasamy, 2017; Leki & Carson, 1994). This issue is particularly evident among foreign students majoring in fields in which the medium of instruction is a language other than their mother tongue. Examples of such students are Iranian students studying in universities in Malaysia. Most Iranian students equate language learning in general, and writing skills in particular, with learning an adequate knowledge of syntax (Baleghizadeh & Golbin, 2010). Consequently, vocabulary is assigned a trivial role. Unfortunately, the lack of vocabulary knowledge has disadvantaged many Iranians studying in higher education (Moghadam, 2012; Shareie, 2001). In a preliminary survey among 20 lecturers from the faculties of education, science, and computer science in a university in Malaysia, (Universiti Teknologi Malaysia – UTM) conducted by the researcher as part of the process of problem identification for this study, the majority of the respondents were dissatisfied with the academic writing ability of their international postgraduate students. They believed that these students, do not generally possess the required academic writing skills to meet the demands of their studies.

## 2. Theoretical framework

Vocabulary knowledge tends to be tightly associated with language proficiency. Studies on vocabulary knowledge reveal the fact that vocabulary is an underlying principle that enables learners to be successful in learning a second language (Laufer & Nation, 2001). On the importance of vocabulary size in language proficiency, Meara (2003) depicts that "all other things being equal, learners with big vocabularies are more proficient in a wide range of language skills than learners with smaller vocabularies" (p. 37). More specifically, some researchers believe in the role of vocabulary size as a predictor for certain language skills (Milton, 2009). Thus, according to Laufer (2010), Meara (2010), and Schmitt (2010), if the correlation between vocabulary size and language skills is significant, a test of vocabulary size can be a good instrument for measuring language skills.

A consistent and reliable correlation between vocabulary size and reading comprehension has been established in many studies (Baleghizadeh & Golbin, 2010; Llach & Gallego, 2009; Nation, 2011). Qian (2002), for example, correlated scores collected from Vocabulary Levels Test (VLT) with the reading comprehension of the TOEFL test administered to 44 Korean and 33 Chinese speakers learning English. The findings of his study figured an even stronger correlation (i.e., 0.78) than that of Laufer (1992) which was .50.

Vocabulary size has been also correlated with writing skill and determined significant correlations. Llach and Gallego (2009) in a study examined the relationship between receptive vocabulary size and written skills (i.e., reading & writing) with two vocabulary size tests (i.e., Word Test & Vocabulary Levels Test), a reading comprehension test and a timed composition, which administered to 274 primary schools Spanish EFL learners. The findings reflected a significant correlation between receptive vocabulary size and reading comprehension (.155 & .156 respectively). However, the results revealed that receptive vocabulary knowledge plays a trivial role in the quality of writing and the estimated correlation was not very strong.

Contrary to written skills, very few studies measured the relationship between vocabulary size and oral skills. Staehr (2008) correlated students' grades on listening, reading, and writing papers with their scores in vocabulary size tests. Similarly, the results suggested a low correlation for listening (0.69) and a high one for reading (0.83) and writing (0.73) skills.

Vocabulary size has been also correlated with speaking skills. "Spoken communication has access to gestures and contextual information which written text usually lacks" (Milton, 2009, p. 177), so it is different from other skills. The results of previous studies suggest that oral skills (i.e., listening & speaking) are not firmly affiliated with vocabulary knowledge.

However, since the focus of this study was mainly on the relationship between vocabulary size and academic writing, and also finding a threshold of vocabulary to predict academic writing skill, some studies on this particular issue are viewed. Academic writing is a type of writing usually used in higher levels of education and for academic purposes. Although interest in academic writing research has increased during the last two decades, the number of studies in this field is not very noticeable. The results of available studies investigating the relationship between vocabulary knowledge and academic writing demonstrate that word knowledge may be a good predictor of academic performance (Laufer & Nation, 1995; Zhu, 2004).

## 2.1. Studies on the threshold level of vocabulary

The vocabulary threshold is the boundary between having and not having enough vocabulary knowledge for language proficiency (Nation, 2007; Tran, et al, 2020). In other words, it is the number of words an L2 learner needs to know in order to be considered as a proficient language user in different language skills. Some studies investigated the relationship between vocabulary size and L2 learners' performance in proficiency tests such as IELTS and Cambridge First Certificate in English (FCE). The findings of these studies have introduced different vocabulary thresholds. Staehr (2008), for example, suggested that at least knowledge of 2000-2500 vocabulary is required for learners to get a score of 5 or better on IELTS speaking and listening sub-skills. However, it is not obvious in his data whether significantly higher vocabulary scores contribute to higher IELTS grades. Moreover, he did not suggest such a threshold for IELTS writing. The present study, first, tried to investigate a correlation between IELTS writing and vocabulary tests. It, then, investigated a vocabulary threshold for academic writing by proposing the two following questions:

1. Is there a significant correlation between productive vocabulary size and academic writing skills?
2. What is the threshold level of vocabulary to predict L2 academic writing skills?

## 3. Methodology

### 3.1. Participants

131 International postgraduate students who obtained the minimum academic (not general) IELTS score essential to get Universiti Teknologi Malaysia (UTM)'s acceptance to enroll in a postgraduate program (i.e., 6 IELTS band score and above) were participants of this study. Due to the validity extension of the IELTS score, only students with IELTS less than 2 years from the date of issuance of their results were chosen as the participants. These students also were chosen randomly from among international students enrolled in postgraduate programs in UTM whether male or female, Ph.D. or Master.

### 3.2. Instruments

Three vocabulary measures were used in this study:

*Productive Vocabulary Levels Test (PVLТ)* measures the size of L2 learners' productive word knowledge. In PVLТ format, half of the word is presented and students complete the word. In fact, they produce the word as they do in writing and speaking skills. The format of PVLТ is as follows:

I'm glad we had this opp\_\_\_\_\_ to talk.

Five different frequency levels are included in the test (i.e., 1-2000, 2000-3000, 3000-5000, University Word List, and 5000-10,000-word level). Each includes eighteen items which makes a total of 90 questions. In previous studies using PVLТ, answers were scored as correct ones if students wrote the correct form of the word considering part of speech, even with some mistakes in grammar or spelling. However, since the purpose of the current study was to measure academic writing ability which is a necessary skill for higher levels of education, the way of scoring was different from the previous research; only those responses which were correct in part of speech, spelling (either American or British spelling), and grammar were considered as correct answers.

*Word Associates Test (WAT)* prepared by Read (1998), designed to measure the depth of vocabulary knowledge by evaluating three important vocabulary components: synonymy, polysemy, and collocation. WAT consists of 40 items. Each

item includes a stimulus word, which is an adjective, and two boxes, each containing four words. The format of WAT is as follows:

### Sudden

Beautiful quick surprising thirsty    Change doctor noise school

For example, in the above table the underlined words are the correct answer. The words ‘*quick*’ and ‘*surprising*’ from the left box are synonymous with the word ‘*sudden*’ and the words ‘*change*’ and ‘*noise*’ from the right box are the collocations (i.e., noun(s) that may come after an adjective in a sentence) of the word ‘*sudden*’. This instruction reduces the opportunity of guessing.

The test taker can choose just four words from two boxes as the collocations of the stimulus word. For example, in the above table, the underlined words are the correct answer. The total score is a number out of 160.

*Lex30* designed by Meara and Fitzpatrick (2000) to measure L2 learners' productive vocabulary size includes 30 stimulus words. All the stimulus words are highly frequent words taken from Nation's (1983) first 1000 wordlist. Any response outside Nation's first 1000-word list is considered as infrequent words. Generally, there is no predetermined answer for stimulus words. The score is a number out of 120. Any mistakes in part of speech or spelling or grammar were considered as the wrong answer. The reliability of *Lex30* is 0.84 estimated by the designers.

In this study, the above three tests were gathered in one test called Vocabulary Test including three parts; each vocabulary test i.e., PVL, WAT, and *Lex30* was considered as one part of the test. The reliability of the new test i.e., Vocabulary Test was .88.

### 3.3. Procedures and research design

Based on the proposed questions, the vocabulary test was first given to participants to measure their vocabulary size. Then, the results of each part of the test were scored and checked for normality using probability plots. The result illustrated that the data were normal and quite suitable for data analysis. In the next stage, the scores obtained from each test were separately correlated with students' scores in the IELTS writing section. Further, Multiple Regression was used to investigate whether participants' score on three tests of vocabulary may predict their writing scores in IELTS.

Afterward, participants' composite scores from the Vocabulary Test including the three tests (i.e., PVL, WAT, & *Lex30*) was correlated with their IELTS writing score. Moreover, a Linear Regression analysis revealed if participants' composite score may predict their writing scores in IELTS.

To find the threshold, the first step was to identify which score students are required in IELTS writing to be considered as competent in academic writing. Based on IELTS descriptors, band score 7 is considered the Good User of language showing that the language learner has operational command of the language, though with occasional inaccuracies, in some situations and generally handles complex language well. This band score is in parallel to the aim of this study i.e., evaluating the writing ability of postgraduate students who are expected to be competent in their language skills, especially writing.

In the second step, since PVL measures the size of productive vocabulary, it was used as the only instrument in this part; the relationship between IELTS writing score and different frequency levels of PVL was shown using simple scatter plots. Then, the participants' mean score at each frequency level of PVL was calculated. Finally, the number of productive vocabularies was estimated for each frequency level of PVL using this formula: (participants' mean score × total words at the level of frequency ÷ 18 (Zimmerman, 2004).

## 4. Result

### 4.1. Result of correlation study

The results of the two-tailed Pearson Product Moment correlation coefficient analysis are summarized in the following table:

**Table 1.** Correlations coefficients between variables

Components of Vocabulary Size	Components of Academic Writing
	IELTS writing score
PVL	.943**
WAT	.938**
Lex30	.796**

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

From the table, it is clear that relationships among the variables were statistically significant. PVLT ( $r=.943$ ), WAT ( $r=.938$ ), and Lex30 ( $r=.796$ ) correlated significantly with students' IELTS writing scores.

## 4.2. Results of multiple regressions

To see if vocabulary tests were able to predict students' scores in IELTS writing, all the variables were analyzed using Multiple Regression (table 2).

**Table 2.** Results of Multiple Regressions

Model	Standardized Coefficients Beta	R	R Square
(Constant)			
Productive Vocabulary Test	.447	.970 <sup>a</sup>	.942
Word Associates Test	.436		
Lex30	.140		

a. predictors: (constant), LEX30, WAT, PVLT

b. dependent variable: IELTS writing score

The table demonstrates the variance of IELTS writing scores from the square of correlation coefficients, i.e. Beta scores. PVLT, for instance, holds the largest variance in IELTS writing (.44) meaning that almost 44% of students' variation in IELTS writing score is predictable by PVLT. WAT possesses the second-largest variation in IELTS writing scores (.436). However, Lex30 could predict a smaller portion of IELTS writing score variation (.140) than the two other tests. Evidently, the results of Multiple Regression revealed that totally a high portion (.97) of the variance of students' IELTS writing scores was predicted by the three tests.

The above-mentioned analyses showed that not only there was a significant correlation coefficient between vocabulary size and academic writing skills, but also the vocabulary tests used in this study especially the Productive Vocabulary Levels Test (PVLT) were able to predict students' results in IELTS writing scores quite well.

## 4.3. Result of the composite score

In order to "tap into different facets of lexical knowledge" (Schmitt, 2010, p. 154), multiple measures were used to measure different aspects of vocabulary (i.e., knowledge of form, meaning, and use). Therefore, participants' composite score from the three tests was calculated and then correlated with their score in IELTS writing.

**Table 3.** Correlation Coefficients between Composite score & IELTS writing score

	IELTS Writing
Composite Score	.965**
N	131

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

From the table, the correlation coefficients between participants' IELTS writing score and their composite scores from the tests were highly significant ( $r=.965$ ,  $P<.01$ ) signifying that any variance in IELTS writing scores relates to any variance in students' composite scores.

A precise comparison between the correlation coefficients of the composite score and IELTS writing with those obtained from each vocabulary test individually (PVLT, WAT, & Lex30) summarized in Table 4, demonstrates that the correlation coefficients between the composite score obtained from the Vocabulary Test was a bit bigger than the three tests on their own. Moreover, the results of Linear Regression (table 5) revealed that a high portion (.96) of the variance of IELTS writing score was predicted by students' composite scores. It simply means that the composite score is a better predictor of IELTS writing scores than the individual tests.

**Table 4.** Correlation Coefficients between IELTS writing score and PVLТ, WAT, & Lex30

Components of Vocabulary Size	Components of Academic Writing
	IELTS writing score
PVLТ	.943**
WAT	.938**
Lex30	.796**
Composite Score	.965**

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

**Table 5.** Coefficients between Composite score and IELTS writing score

Model	Standardized Coefficients Beta	R	R Square
1 (Constant)			
Composite Score	.965	.965 <sup>a</sup>	.931

a. Dependent Variable: IELTS

b. Predictor: (Constant), Composite score

The results shown in the tables support the idea that investigating vocabulary knowledge from different aspects provides us with a comprehensive picture of L2 learners' word knowledge.

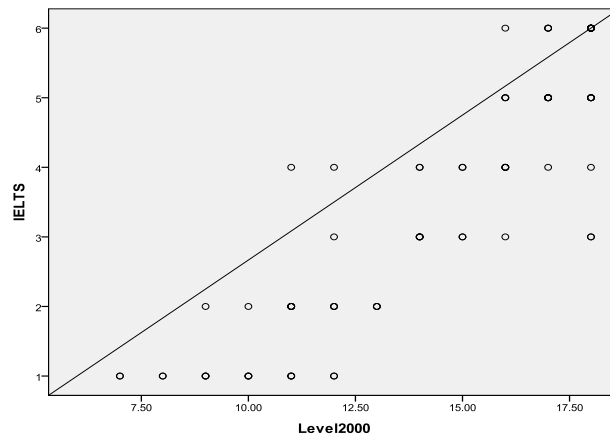
#### 4.4. Results of threshold

To find the threshold, first, the mean of participants' scores at each frequency level of PVLТ was calculated as presented in Table 6. Then, the number of productive vocabularies was estimated for each frequency level of PVLТ using this formula: (participants' mean score × total words at the level of frequency ÷ 18 (Zimmerman, 2004)).

**Table 6.** The mean scores and percentage for each band scores of IELTS writing

IELTS Writing Band Scores	n	PVLТ (90)		Level 2,000		Level 3,000		Level 5,000		Level 10,000		WAT (160)		Lex30 (120)	
		Mean	%	Mean	%	Mean	%	Mean	%	Mean	%	Mean	%	Mean	%
8	7	73.14	81	18.0000	100	13.2857	74	13.1429	73	11.1429	62	141.71	89	115.86	97
7.5	20	69.35	77	17.6500	98	15.3500	85	13.2000	73	6.8000	38	125.95	79	92.70	77
7	22	52.32	58	16.9091	94	11.6364	65	8.6818	48	3.0000	17	120.27	75	93.18	78
6.5	22	48.50	54	15.2727	85	10.0000	56	8.6364	48	3.0000	17	99.55	62	86.82	72
6	21	42.62	47	14.5714	81	8.9048	50	6.7143	37	1.8095	10	91.29	57	73.48	61
5.5	19	30.11	34	11.3158	63	5.9474	33	4.0526	23	.9474	5	69.53	44	69.89	58
5 20 Total	131	23.35	26	9.6500	54	4.6500	26	2.5000	14	.4000	2	53.15	33	60.50	50

Accordingly, the relationship between IELTS writing score (i.e., vertical axis; 1=IELTS band score 5; 2= IELTS band score 5.5; 3= IELTS band score 6; 4= IELTS band score 6.5; 5= IELTS band score 7; 6= IELTS band score 7.5; and 7= IELTS band score 8.) and different frequency levels of PVLТ (i.e., horizontal axis) was identified using simple scatter plots.



**Figure 1.** IELTS scores at 2000 frequency level

The above graph (figure 1) indicates that participants in band 7 of IELTS writing could answer 14 to 18 questions out of 18 at level 2,000 of PVLТ correctly. It means that they knew approximately 1,777 words at this level ( $16 \times 2,000 \div 18 = 1,777$ ). This number was derived by converting the raw scores into an estimated number of words by multiplying the mean score by the related number of frequency levels and dividing by 18 (i.e., the number of items in each frequency level). The result was the total number of words known by participants at that level. Based on this formula, the number of words known by the learners at different levels of frequency is calculated and summarized in Table 7.

**Table 7.** Estimated Productive Vocabulary Size for different IELTS writing band scores

IELTS Writing Band Scores	n	Word Frequency Level			
		2000	3000	5000	10000
5	20	1,000	666	555	0
5.5	19	1,222	833	1,111	522
6	21	1,555	1,333	1,666	555
6.5	22	1,666	1,666	2,222	1,666
7	22	1,777	1,833	2,222	1,666
7.5	20	1,888	2,500	3,611	3,333
8	7	2,000	2,166	3,611	6,111

The table shows that the number of words known by participants increases slightly through different IELTS writing band scores. It means that those participants who knew more words could get a higher score in academic writing of IELTS.

However, the point is that based on Nation's descriptor of his test which is provided in <http://www.lex tutor.ca/tests/levels/productive/10ka.html>, and summarized in table 8., the percentage of the mean score of participants at band score 7 of IELTS writing for level 2000 is 94% that is remarkably above the cut of score for that level i.e. 83% identified by Nation (2001). While for other levels the estimated percentage is below the cut of score. It means that participants with a band score of 7 in IELTS writing could pass the first level of frequency i.e., 2000 in PVLТ.

**Table 8.** Descriptors for frequency levels of PVLТ

Frequency Levels of PVLТ	Threshold Level (accepted percentage)	Percentages obtained in this Study for participants with 7 in IELTS Writing
Level 2000	83%	94%
Level 3000	83%	65%
Level 5000	83%	48%
Level 10,000	80%	17%

To sum up, looking at the findings of this study it could be concluded that a full knowledge of 2,000 words would lead to getting 7 in IELTS writing. In other words, 2000 productive words are the threshold level of vocabulary for academic writing skills. It means that if an L2 learner wishes to be competent in academic writing skills, s/he should have full knowledge of 2000 productive words. Therefore, the answer to the second question of this study was also identified.

## 5. Discussion

In light of the results of the Pearson correlation analysis which was related to the first question proposed in this paper, a significant and strong correlation coefficient was established between students' writing scores and their scores in tests of vocabulary (i.e., Productive Vocabulary Levels Tests, Word Associates Test, & Lex30) ( $r=.94, .93, \& .79$  respectively;  $p<.05$ ). The results indicate that generally there is a highly significant relationship between productive vocabulary size and academic writing skills. These results are analogous to the prior findings by Zimmerman (2004) who found a high correlation ( $r = .60$ ) between students' scores on the Productive Vocabulary Levels Test (PVL) and their scores on the writing test that was considerably lower than the one of this study (.96). The reason for this big difference may lie in this fact that receptive vocabulary is not well related to academic writing. Thus, we might safely argue that productive vocabulary size is more correlated with writing skills rather than receptive vocabulary size.

Multiple Regression analysis revealed that there was a statistically significant interaction between IELTS writing scores and vocabulary tests ( $F=68, p>0.001$ ). Therefore, it appears that PVL, WAT, and Lext30 may predict grades in the IELTS writing examination well, among which PVL specified the most variation of participants' academic writing ( $t=8.729$ ). The findings of this study which are in line with a study carried out by Meara and Milton (2003) in which vocabulary size has been proved to predict participants' performance in English language Tests, highlighted the significant relationship between vocabulary knowledge and L2 learners' performance in English skills.

As discussed earlier, multiple measures were used to measure special aspects of word knowledge. PVL and Lex30 measured the size of productive vocabulary, while WAT evaluated knowledge of the depth of vocabulary. This study has succeeded in showing that depth, as well as the size of vocabulary knowledge, is well related to academic writing ability. The composite score, which was the participants' total score from the three tests i.e., PVL, WAT, and Lex30, also showed a significantly high correlation coefficient (.96) with their IELTS. However, in other studies using multiple instruments (Akbarian, 2010; Zimmerman, 2004; & Webb, 2008), no attempt was made to calculate the composite score.

The highly significant correlation coefficients established in this study support the idea that multiple tests should be used to provide us with a more comprehensive picture of students' vocabulary knowledge. Therefore, the current study was probably unique in using the composite score to measure L2 learners' word knowledge. Further, the significant correlation coefficients between participants' composite score and their IELTS writing scores (.96) indicate that academic writing skills are highly interrelated with the size and depth of L2 learners' word knowledge. However, the significant correlation coefficients between vocabulary tests and participants' academic writing should not blind us to the fact that academic writing is a complex skill that may not be measured by using vocabulary tests. Being good writers, L2 learners should know other basic skills of writing such as prewriting, planning, using transitions, producing a clear thesis statement and topic sentences, developing paragraphs, and controlling punctuation. Moreover, knowledge of the conventions of academic writing, thorough knowledge of the discipline, the type of research/topic written on, and the nature of the field itself interact in a complex way with language proficiency, especially on the quality of academic writing. Therefore, using vocabulary tests to measure academic writing may challenge only one dimension of academic writing i.e., vocabulary.

The results of Scatter Plots and Descriptive Analysis revealed a threshold level for vocabulary knowledge which was the focus of the second question of this study. The scatter plots indicate that participants who obtained a band score of 7 in IELTS writing which was identified as the competent level of academic writing skill based on the IELTS band score descriptor were competent in the first 2000 level of frequency. In other words, full mastery of 2000 words is most likely required for an L2 learner to get a band score of 7 in IELTS writing. A vocabulary size below this threshold i.e., 2000 words is probably insufficient for competency in academic writing skills which was much less than that obtained in Meara and Milton's (2003) who estimated a threshold of 3500 words for being accepted in Cambridge First Certificate in English (FCE). Two main reasons tended to strongly influence the vocabulary threshold identified in this study. First, the threshold identified in this study was based on the number of learners' productive word knowledge since it was believed that productive vocabulary is more related to academic writing skills rather than receptive vocabulary and is usually smaller than receptive one. A second reason is based on the idea that academic writing is a skill for higher levels of education in which a higher level of English literacy is demanded, a very strict scoring system which was used to evaluate participants' full word knowledge in academia. However, it seems that the result is still consistent with Staehr's (2008) who reports a threshold level of 2000 words that must be crossed if learners are to gain an average score or above in writing.

Eventually, such co-relational studies may only give a small piece of a larger picture; an indication of the extent of vocabulary knowledge needed for academic writing. Future research might include other aspects involved in academic writing such as individual variation or learners' motivation.



## 6. Conclusion

An appropriate final word would seem to be that vocabulary is highly related to communicative skills such as speaking and writing. In a bigger picture, the more vocabulary L2 learners know, the better they may operate in English. With these considerations, successful learners are likely those who manage their vocabulary acquisition by increasing the size of their vocabulary knowledge as well as learning other aspects of vocabulary such as vocabulary depth, word association, and collocation. On the part of the learner, a deliberate effort is required to acquire a huge amount of vocabulary over an extended time. However, the volume of acquired vocabulary may vary from one context to another. For example, in a context in which English is spoken as a second language learners' exposure to English is much greater than in a context in which English is considered as a foreign language. Apparently, researchers should be careful when generalizing the findings of vocabulary measurements from one group of participants to a wider population. On the part of teachers, it is highly recommended that they, first, include teaching vocabulary explicitly in their teaching syllabus. Secondly, it would be a good practice if they teach students about other aspects of vocabulary such as association and collocation when they present new words. On the whole, the findings of this study highlight the importance of vocabulary knowledge in learning a second language.

Three limitations were identified in this study. First, the number of participants was not equal and sufficient for each of the IELTS writing band scores. For example, there were only 8 participants for band score 8, while for other band scores, this number was between 19 and 22 people. Moreover, no participants were identified with IELTS writing band scores of 8.5 and 9. The limited number of participants in each IELTS writing band score may reduce the reliability of the results.

Considering the participants of the study, a second limitation is that the majority were Iranian students who had learned English as a foreign language. Consequently, the generalization of the results of this study to second language learners should be done with caution.

The third limitation was that among 131 participants who participated in this study only 45 of them provided their supervisor's (lecturer's) name to assess their actual ability in academic writing. Hence, getting 7 in IELTS writing which was agreed to be the boundary for being competent in academic writing by the lecturers may vary in another context with a higher number of lecturers' assessments. Moreover, reliance on lecturers' assessments to evaluate their students' academic writing competency based on brief descriptors seemed to be insufficient to assess the students' genuine academic writing ability.

Some ideas for future research were elicited from this study. First of all, as mentioned before the number of participants was not sufficient for each of the IELTS writing band scores. Therefore, future research could be carried out among different International Universities to include more participants with IELTS different band scores. However, it was not viable for the present study.

Second, future research may carry out the tests in different faculties separately. The underlying idea is that different faculties have different disciplines for academic writing. For example, in the Faculty of Education, language performance expectations are higher than in the Faculty of Engineering or Science.

Finally, using IELTS and lecturers' assessments as instruments to measure students' academic writing simply could be replaced by another tool such as free writing. Asking students to write could best measure their actual academic writing skills if there was a standard criterion for scoring. Thus, future studies may make use of free writing to measure students' academic writing skills. However, asking students to sit in a long session of writing is time-consuming and was not feasible in this study.

## 7. References

- Afshari, S., & Tavakoli, M. (2017). The relationship between depth and breadth of vocabulary knowledge and Iranian EFL learners' listening comprehension. *International Journal of Research Studies in Language Learning*, 3(6), 13-24. <https://doi.org/10.5861/ijrsl.2016.1438>
- Akbarian, I. (2010). The relationship between vocabulary size and depth for ESP/EAP learners. *Science Direct*, 391-401. <https://doi.org/10.1016/j.system.2010.06.013>
- Alderson, J. C. (2007). Judging the frequency of English words. *Applied Linguistics*, 28 (3) 383-409. <https://doi.org/10.1093/applin/amm024>
- Aliabadi, H., Tazik, H., & Azizi, Z. (2022). Exploring the role of language knowledge and background knowledge in reading comprehension of specific-purpose tests in higher education. *Language Testing in Asia*, 12 (1). <https://doi.org/10.1186/s40468-022-00198-x>

- ALSager, R., & Milton, J. (2016). Investigating the Relationship between Vocabulary Knowledge and Academic Success of Arabic Undergraduate Learners in Swansea University. *LIF – Language in Focus Journal*, 2, <https://doi.org/10.1515/lifijjal-2016-0010>
- Anderson, R. C., & Freebody, P. (1981). Vocabulary knowledge. In J. T. Guthrie (Ed.). *Comprehension and Teaching: Research Reviews* (pp. 77-117).
- Baleghizadeh, S., & Golbin, M. (2010). The effect of vocabulary size on reading comprehension of Iranian EFL learners. *Linguistic and literacy broad research and innovation*, 1, (2), 33-47.
- Clarke, M. (1979). Reading in Spanish and English: Evidence from adult ESL students. *Language Learning*, 29, 121-150. <https://doi.org/10.1111/j.1467-1770.1979.tb01055.x>
- Cui, Y. (2008). *L2 Proficiency and L2 Reading: Consolidating the Linguistic Threshold Hypothesis*. University of Victoria.
- Daller, H., Milton, J., & Treffers-Daller, J. (Eds) (2007). *Modelling and Assessing Vocabulary Knowledge*. Cambridge University Press.
- Elgort, I., & Nation, I. S. P. (2010). Vocabulary Learning in a Second Language: Familiar Answers to New Questions. In P. Seedhouse, S. Walsh & C. Jenks (Eds). *Conceptualising 'Learning' in Applied Linguistics* (pp. 89-104). Palgrave Macmillan. [https://doi.org/10.1057/9780230289772\\_6](https://doi.org/10.1057/9780230289772_6)
- Ellis, R. (1994). Factors in the incidental acquisition of second language vocabulary from oral input: a review essay. *Applied Language Learning*, 5(1), 1-32. <https://doi.org/10.1075/sibil.17.06ell>
- Ellis, R. (1995). Modified oral input and the acquisition of word meanings. *Applied Linguistics*, 16(4), 409-441. <https://doi.org/10.1093/applin/16.4.409>
- Haastrup, K., & Henriksen, B. (2000). Vocabulary acquisition: acquiring depth of knowledge through network building. *International Journal of Applied Linguistics*, 10(2), 221-240. <https://doi.org/10.1111/j.1473-4192.2000.tb00149.x>
- Hirsh, D. (1992). *The vocabulary demands and vocabulary learning opportunities in short novels*. Victoria University of Wellington, Wellington.
- Hulstijn, J. (2003). Incidental and intentional learning. In C. Doughty & M. Long (Eds.), *Handbook of Second Language Acquisition* (pp. 349-381). Blackwell. <https://doi.org/10.1002/9780470756492.ch12>
- Laufer, B. (1992). How much lexis is necessary for reading comprehension? In P. J. L. Arnaud & H. Bejoint (Eds.), *Vocabulary and Applied Linguistics* (pp. 126-132). London: Macmillan. [https://doi.org/10.1007/978-1-349-12396-4\\_12](https://doi.org/10.1007/978-1-349-12396-4_12)
- Laufer, B., & Nation, P. (1995). Vocabulary size and use: lexical richness in L2 written production. *Applied Linguistics*, 16(3), 307-322. <https://doi.org/10.1093/applin/16.3.307>
- Laufer, B., & Nation, P. (2001). Passive vocabulary size and speed of meaning recognition: are they related? *EUROSLA Yearbook*, 1, 7-28. <https://doi.org/10.1075/eurosla.1.05lau>
- Laufer, B., & Paribakht, T. S. (1998). The relationship between passive and active vocabularies: effects of language learning context. *Language Learning*, 48(3), 365- 391. <https://doi.org/10.1111/0023-8333.00046>
- Laufer, B., & Ravenhorst-Kalovski, G. C. (2010). Lexical threshold revisited: Lexical text coverage, learners' vocabulary size and reading comprehension. *Reading in a Foreign Language*, 22(1), 15-30.
- Laufer, B., & Yano, Y. (2001). Understanding unfamiliar words in a text: Do learners understand how much they don't understand? *Reading in a Foreign Language*, 13(2), 549-566.
- Lee, C. & Rethinasamy, S. (2017). Breadth and depth of academic vocabulary knowledge: Assessing their roles in academic reading and writing of English language learners. *MJAL*, 9 (3), 196-224.

- Lee, S. H., & Muncie, J. (2006). From receptive to productive: Improving ESL learners' use of vocabulary in a post reading composition task. *TESOL Quarterly*, 40(2), 295- 320. <https://doi.org/10.2307/40264524>
- Leki, I., & Carson, J. G. (1994). Students' perceptions of EAP writing instruction and writing needs across the disciplines. *TESOL Quarterly*, 28(1), 81-101. <https://doi.org/10.2307/3587199>
- Li, J. & Schmitt, N. (2009). The acquisition of lexical phrases in academic writing: A longitudinal case study. *Journal of Second Language Writing*, 18, 85–102. <https://doi.org/10.1016/j.jslw.2009.02.001>
- Llach, M. P. A. & Gallego, M. T. (2009). Examining the Relationship between Receptive Vocabulary Size and Written Skills of Primary School Learners. *Journal of the Spanish Association of Anglo-American Studies*, 31 (1), 129–147.
- Meara, P. (2000). Lex30: an improved method of assessing productive vocabulary in an L2. *System*, 28, 19-30. [https://doi.org/10.1016/S0346-251X\(99\)00058-5](https://doi.org/10.1016/S0346-251X(99)00058-5)
- Meara, P. (2010). The relationship between L2 vocabulary knowledge and L2 vocabulary use. In Macaro, E. (Ed). *The continuum companion to second language acquisition* (pp. 179-193). Continuum.
- Meara, P., & Alcoy, J. C. O. (2010). Words as species: An alternative approach to estimating productive vocabulary size. *Reading in a Foreign Language*, 22(1), 222- 236.
- Meara, P., & Fitzpatrick, T. (2000). Lex30: An improved method of assessing productive vocabulary in an L2. *System*, 28(1), 19-30. [https://doi.org/10.1016/S0346-251X\(99\)00058-5](https://doi.org/10.1016/S0346-251X(99)00058-5)
- Meara, P. & Milton, J. (2003). *X\_Lex, The Swansea Levels Test*. Newbury Express.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Multilingual Matters. <https://doi.org/10.21832/9781847692092>
- Moghadam, S. (2012). *The relationship between Breadth and depth of vocabulary knowledge And Reading Comprehension*. Unpublished Doctoral Dissertation, Universiti Teknologi Malaysia.
- Mohan, B. A. & Lo, W.A.Y. (1985). Academic Writing and Chinese Students: Transfer and Development Factors. *TESOL Quarterly*, 19(3), pp. 515-533.
- Morris, L., & Cobb, T. (2004). Vocabulary profiles as predictors of the academic performance of Teaching English as a Second Language trainees. *System*, 32, 75-87. <https://doi.org/10.1016/j.system.2003.05.001>
- Nasir, N., Ab Manan, N. & Azizan, N. (2017). Examining the relationship between vocabulary knowledge and General English Language Proficiency. *Esteem Journal of Social Sciences and Humanities 1*, 15-22.
- Nation, I. S. P. (1983). Testing and teaching vocabulary. *Guidelines*, 5(1), 12-25. <https://doi.org/10.1002/9781444315783.ch28>
- Nation, I. S. P. (2001). *Learning Vocabulary in Another Language*. Cambridge University Press.
- Nation, I. S. P. (2007). Vocabulary learning through experience tasks. *Language Forum*, 33(2), 33-43. [https://doi.org/10.1057/9781137449405\\_2](https://doi.org/10.1057/9781137449405_2)
- Qian, D. (2002). Investigating the relationship between vocabulary knowledge and academic reading performance: an assessment perspective. *Language Learning*, 52(3), 513-536. <https://doi.org/10.1111/1467-9922.00193>
- Read, J. (2007). *Review of "Focus on Vocabulary"*. *New Zealand Studies in Applied Linguistics*, 13(1), 101-103.
- Schmitt, N. (2010). *Researching Vocabulary: A Vocabulary Research Manual*. Palgrave Macmillan.
- Schmitt, N., Schmitt, D. & Clapham, C. (2001). Developing and exploring the behavior of two new versions of the Vocabulary Levels Test. *Language Testing*, 18, 55-88. <https://doi.org/10.1177/026553220101800103>

- Shareie, M., Salehi, S., & Fazeli, M. (2001). Investigating the essential values among university students. *Social Sciences*, 63-88.
- Staehr, L. S. (2008) Vocabulary size and the skills of listening, reading and writing. *Language Learning Journal*, 36 (2), 139-152. <https://doi.org/10.1080/09571730802389975>
- Tran, A. H., Tremblay, K. A., & Binder, K. S. (2020). The factor structure of vocabulary: An investigation of breadth and depth of adults with low literacy skills. *J Psycholinguist Res* 49, 335–350. <https://doi.org/10.1007/s10936-020-09694-8>
- Vermeer, A. (2001). Breadth and depth of vocabulary in relation to L1/L2 acquisition and frequency of input. *Applied Psycholinguistics*, 22, 217–234. <https://doi.org/10.1017/S0142716401002041>
- Webb, S. (2008). Receptive and productive vocabulary sizes of L2 learners. *Studies in Second Language Acquisition*, 30, 79-95. <https://doi.org/10.1017/S0272263108080042>
- Wilkins, D. A. (1972). *Linguistics in Language Teaching*. Arnold. doi: <https://doi.org/10.1017/S0025100300000682>
- Zhu, W. (2004). Faculty views on the importance of writing, the nature of academic writing, and teaching and responding to writing in the disciplines. *Journal of Second Language Writing*, 13, 29–48. <https://doi.org/10.1016/j.jslw.2004.04.004>
- Zimmerman, K. J. (2004). *The role of vocabulary size in assessing second language proficiency*. Unpublished Doctoral Dissertation, Brigham Young University.



<https://doi.org/10.22077/ali.2023.6612.1019>

# Indirect Complaint as an Act of Rapport-Inspiring Speech Behavior: The Case of Iranian Students in the University Context

Esmaeel Ali Salimi<sup>1\*</sup>  
 Meysam Khazae Kouhpar<sup>2</sup>

<sup>1</sup>Associate Professor in Applied Linguistics, Department of English Language and Literature, Faculty of Persian Literature and Foreign Languages, Allameh Tabataba'i University, Tehran, Iran

<sup>2</sup>PhD Candidate in Teaching English as a Foreign Language, Department of English Language and Literature, Faculty of Persian Literature and Foreign Languages, Allameh Tabataba'i University, Tehran, Iran

## ARTICLE HISTORY

Received: 26 June 2023  
 Revised: 27 July 2023  
 Accepted: 23 August 2023  
 Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [esalimi@atu.ac.ir](mailto:esalimi@atu.ac.ir)

## ABSTRACT

Indirect complaints (ICs) most often serve as rapport-inspiring speech behavior, whose affiliation by the recipient of the complaint preserves solidarity among peers. The purpose of the present paper was to investigate the themes and response strategies used by Iranian students as an act of IC in the university context. Drawing on a mixed methods research design, the results of a thorough thematic analysis of the collected conversational data revealed that the themes Iranian students tend to use as an IC are not independent of their gender. In addition, two new IC response strategies were identified, namely attribution and admonishment. Hence, in order to account for all of the strategies used by Iranian students to respond to ICs, the existing categorizations are to be expanded. It was also discovered that only through explicit or embedded commiseration is there more potential for longer interactions among participants of conversations. The present study sheds some light on the pragmatics of the Persian language.

**KEYWORDS:** Indirect complaint; IC themes; IC response strategies; Rapport-inspiring speech behavior; Speech act of complaint

## 1. Introduction

University life can be very stressful, especially for students not ready to survive the major upheaval of transiting to a new educational environment (Bedewy & Gabriel, 2015; Hathaway et al., 2023). The situation can be aggravated if this transition involves moving to big cities where the new lifestyles and cultural norms of the local people might stand in stark contrast to students' own ways of life (Gundogan, 2023). Research conducted in universities exploring students' accounts (in particular those provided by working-class students) of their experiences in higher education has revealed a myriad of educational inequalities (Leathwood & O'Connell, 2003; Ulriksen et al., 2017) that have frustrated students' aspirations to progress to higher education. In this regard, Archer et al. (2003) contend that a great many young people from working-class status are loath to go on to higher education as their aspirations and self-esteem have been inauspiciously affected by various kinds of

discrimination in society, the great source of which is probably the practices of universities themselves. Such dissatisfaction with educational services in higher education is, to a great extent, likely to be expressed in the form of ‘complaining’ to friends and other peers (Jääskä & Aaltonen, 2022). The present study intends to shed some light on this commonplace social skill.

Merriam-Webster’s Collegiate Dictionary defines the noun “complaint” as “expression of grief, pain, or dissatisfaction.” However, the elusive definition of “a complaint” as an interactional activity is difficult to provide (Ruusuvuori & Lindfors, 2009). A more comprehensive definition of the term is to account for the notion of individuals’ public self-image or self and the way it relates to the negative stance people tend to take while participating in the activity of complaining. The term “face,” as defined by Goffman (1967) refers to “the positive social value a person effectively claims for himself by the line others assume he has taken during a particular contact” (p. 5). This concept covers various aspects of ourselves and is strongly related to our perception of what it means to be embarrassed, respected, appreciated, and accepted or refused to name but a few (House & Kádár, 2023). Since our face is more likely to be threatened by our interlocutors as an act of self-defense if their face is not saved, we often tend to make sure everyone’s face is maintained (and even enhanced) in our daily interactions (Brown & Levinson, 1987). For instance, Tseng and Chen (2022), analyzing some sports data, have discovered that individuals whose social duty necessitates reaching communicative concord in certain communicative contexts that entail tension are more likely to utilize mutual face-maintaining acts in order to avoid any conflict. This clearly shows that “everyone’s face depends on everyone else’s being maintained” (Brown & Levinson, p. 61).

According to Brown and Levinson (1987), “certain kinds of acts intrinsically threaten face” (p. 65). In other words, there are face-threatening acts (FTAs) that by nature, stand in contrast to the ‘face wants’ of either interactant. Complaining is regarded as one of the main face-threatening acts (some more examples of FTAs include criticism, insult, and expression of disapproval). As Brown and Levinson (1987) argued, complaining shows that the speaker has a negative evaluation of a certain aspect of the hearer’s positive face and that he does not care about the addressee’s wants. However, such a perspective toward face in politeness theory has been extensively criticized (see, for example, Izadi, 2017; Spencer-Oatey, 2007). For one thing, not all speech acts are face-threatening in need of being mitigated (Arundale, 2006). On the other, in their conceptualization of the face, Brown and Levinson (1987) paid little, if any, attention to the social dimension of the face (Tseng & Chen, 2022).

In the present paper, the speech act of complaint refers to an activity (in the form of an utterance) that communicates displeasure to an action or some target that has unfavorably treated the complainant (Schegloff, 2005). Here, the specific context where complaining occurs is the university context, and the complainants are its students. What is more, in order to account for the social dimension of the face, university students’ utterances and response strategies are examined as a set of social skills in light of Goffman’s (1967) conceptualization of the face.

## 2. Literature Review

### 2.1. The speech act of complaint

Initial research on the use of complaints has centered on the identifiable sequences of complaints, in which the beginning and end are readily recognizable (e.g., Drew, 1998). Espousing complaints’ identifiability through explicit devices such as “idiomatic expressions” (Drew & Holt, 1988), “negative observations” (Schegloff, 1988), and “extreme case formulations” (Pomerantz, 1986), more recent studies have dealt with issues such as the way the activity of complaint develops (Decock & Depraetere, 2018; Traverso, 2009) and whether and how the recipients of complaints affiliate or disaffiliate with an instance of complaint (Drew & Walker, 2009; Rodriguez, 2022).

The speech act of complaint has also been explored in the fields of interlanguage as well as cross-cultural pragmatics (refer to Fogal et al. (2018) for a summary of recent advances in the field of speech acts). For instance, Tatsuki (2000) compared and contrasted the use of complaints by Japanese students in both Japanese and English. Her findings, which were the result of a psychological test, revealed that Japanese students used different types of aggression in English and Japanese when responding to frustration or stress. This, she further argued, might result in more severe instances of complaints in English than in Japanese. Finally, she suggested that attempts should be made to make Japanese learners of English more aware of different levels of politeness in English.

In order to perform an appropriate complaint, which is typically associated with an FTA, one needs to be aware of the different components of this speech act. This is crucial, especially for non-native speakers of any particular language (Boxer & Pickering, 1995), for it seems so easy to fall into the trap of mistaking criticism for complaint. As an example, Murphy and Neu (1996) investigated the appropriacy of complaint speech act set by Korean non-native speakers of English through the perception of American native speakers of English. Accordingly, they conducted a two-part study exploring the production of complaint by both native and non-native speakers, and further analyzed native speakers’ judgments of non-native speakers’ productions. As a result, their analysis of the collected data uncovered a significant deviation from appropriate codes of complaints by Korean non-native speakers of English, most of whom had produced, in native speakers’ perception, a criticism, instead of a complaint.

## 2.2. Direct complaint versus indirect complaint

Previous studies suggest two types in the categorization of complaints: direct (Dersley & Wootton, 2000, among others) and indirect (see, for example, Boxer, 1996; Drew, 1998). Indirect complaints (ICs) are different from direct complaints (DCs) in that the addressee per se is neither considered responsible for an instance of the offense nor unfavorably evaluated (Boxer, 1996). Rather, IC will be defined here as “the expression of dissatisfaction to an interlocutor about oneself or someone/something that is not present” (Boxer, p. 219). Even though a number of studies have been carried out on the speech act of complaint in different contexts and cultures (e.g., English, Japanese, French, and Italian), there are relatively few studies that have focused on complaints in the Iranian context. What is more, most studies related to the speech act of complaint have either focused on DCs or the complaints themselves rather than on ICs and IC response strategies (Boxer, 1993a). The present study aims at exploring the use of the speech act of complaint by Iranian students in the university context. Its main concern, however, has to do with ICs, as opposed to DCs.

DCs are the result of dissatisfaction with a certain aspect of the complainee that is likely to spawn an argument. In order to avoid an argument, both interactants need to make use of certain strategies to restrict the impact of the complaint. Laforest (2002) investigated the complaint/complaint-response sequence in everyday conversations between four French-speaking Montréal families, consisting of 50 hours of family conversations. The result of this study indicated that entry into an argument is usually the consequence of questioning the value of the complainee’s response. Furthermore, it was suggested that complaining is a useful means of controlling another person’s behavior. Therefore, expression of dissatisfaction is appreciated as long as it does not lead the interaction to a verbal confrontation.

As mentioned earlier, ICs differ from DCs in that the present addressee is not considered responsible for a certain instance of complaint (Boxer, 1996). Generally speaking, it is safe to say that ICs are not designed to threaten the addressee’s face. Instead, ICs often serve as rapport-inspiring speech behavior, intended to preserve peer solidarity (Boxer & Pickering, 1995; Rodriguez, 2022).

## 2.3. Indirect complaint, gender, and building up rapport

Literature concerning the analysis of social interaction indicates that expressing dissatisfaction about an absent party or the very situation in which the complaining occurs serves as a highly potent device through which a closer relationship is forged (e.g., Goodwin, 1990; Rodriguez, 2022). Nevertheless, once the complainant confides in their friends about their discontent or grievance against an absent party, i.e., when an instance of IC occurs, the complainant is at their most vulnerable. This is because the complaint recipient may simply not be interested in further engagement in the activity, as a result of which the complainant’s hope of being paid attention to might be disappointed. Meanwhile, results from several studies pertinent to the function of ICs suggest that ICs in the sequences of troubles-telling in everyday conversation (Jefferson, 1988) tend to encourage affiliation with the complainant more frequently (Boxer, 1993a, 1993b, 1996; Kozlova, 2004).

From a socio-pragmatic angle, gender differences have been considered as an effective variation in politeness literature (e.g., Holmes, 1990, 1995, 2006; Mills 2003), especially in Brown and Levinson’s (1987) framework. As Leech (2014) highlighted studies on gender specification and social parameters can expand our awareness of the two counterparts’ social status.

In the context of Iran, Allami (2006) analyzed different replies to gripping (IC) by Iranian students. His findings indicate that in the Iranian context, ICs are primarily used for commiseration rather than other purposes. In addition, he found that Iranian students feel obliged to commiserate with the complaining speaker in furtherance of keeping solidarity with one another. With regard to gender differences, the remainder of his findings did not support Boxer’s. While Boxer argued that women generally tend to commiserate more with the complaining speaker and that men are mostly inclined to give advice or contradict, Allami challenged this finding contending that both male and female groups pay equal attention to commiseration. This non-face-threatening act, he argued, becomes manifest not only among friends but also among strangers.

As far as politeness and appropriateness of directness are concerned, Iranian culture turns out to be associated with its distinct Islamic values. One key factor that distinguishes Iranian culture from western culture is probably the level of *individualism*. In this regard, Iran, with a score of 41, is believed to be a collectivist society (Hofstede et al., 2010). Drawing on this consideration, one can better justify the reason why, in Persian, ICs and even some direct speech acts underscore loyalty, solidarity and in-group membership (see for instance Eslamirasekh, 1993). After all, in collectivist societies, an instance of offense (and therefore complaining about it) will probably lead to embarrassment and loss of face (Hofstede et al., 2010).

The present study aims to determine the themes and response strategies used by Iranian students as an act of IC in the university context. Besides, this study aims at exploring the relationship, as well as the strength of the association, between Iranian students’ gender and the themes they tend to use as an IC. In doing so, the research attempts to find the answers to the following questions:

1. What themes do Iranian students tend to use in the university context as an indirect complaint?
2. Is there any statistically significant relationship between Iranian students' gender and the themes they tend to use as an indirect complaint?
3. What strategies do Iranian students prefer to use in university context as an indirect complaint response?

### 3. Method

#### 3.1. Corpus

The data consists of 239 IC exchanges in Persian, recorded at Allameh Tabataba'i University, one of the universities of Iran, located in northwestern Tehran. After data collection and an initial rough transcription of the Ics, it turned out that of more than 15500 students in this university, 68 (34 male and 34 female) were recorded while casually conversing with peers in Persian. The students in this study were either undergraduates or postgraduates majoring in Persian literature and foreign languages. These were individuals with whom the researchers had different levels of friendship.

#### 3.2. Procedure

Conversational data from students were collected from spontaneous speech that was either audio-recorded using Audio Recorder Apps, available on various smartphones, or recorded in the form of field notes, immediately after leaving the site (this is indicated in the accompanying examples provided in *Results and discussion* section). Once the data were gathered and an initial rough transcription was prepared, they were transcribed verbatim and annotated using QSR Nvivo software. An important point should be made here. A statement varying in length from a single phrase to a whole paragraph was counted as an IC if a 'pronounced negative stance' (Ruusuvuori & Lindfors, 2009) embedded in any form of 'troubles-telling narrative' (Jefferson, 1988) was adopted toward any possible 'complainable' (Schegloff, 2005), except those directed at the recipient of the complaint (i.e., DCs were not counted). Throughout data collection, the students were audio-recorded or recorded in the form of field notes on various sites within the main university area, some of which include the campus, the classrooms, the self-service area, and the library site.

Almost 80% of the data was audio-recorded. A majority of audio-recording was done with prior consent directly from the participants. Only for a small number of cases (less than 10%), permission was obtained after it was revealed to the students that a recording had been made. The data gathered in the form of field notes were mostly overheard by the researchers without any participation in the conversations taking place. The data were collected over a period of approximately nine months in 2022 and 2023. The total number of hours transcribed equals 27 hours, more or less; however, only the relevant excerpts from the whole corpus with regard to the purpose of the present study were analyzed. Thus, as mentioned earlier, the data consists of audio recordings and filed notes of 239 IC exchanges. These instances were uttered by a vast number of university students (n=68), so the researchers looked into a student population rather than the mere idiosyncrasies of a handful of people.

#### 3.3. Data analysis

Using a mixed methods research design and drawing on the categorizations proposed by (Boxer, 1993b), qualitative data were analyzed by means of thematic analysis. Thematic analysis is one of the most common methods of analyzing qualitative data (Braun & Clarke, 2006), through which it is possible to pinpoint and report recurrent patterns (themes) within the data set. This means the researchers started the analysis first by acquainting themselves with the data in order to identify common patterns among the Ics. This led to certain initial codes based upon bodies of language, such as single phrases and whole paragraphs, that indicated discontent with oneself, the situation, or an absent party, after which relevant themes were identified. Next, the identified themes were reviewed, refined, and named, and finally, a report of the analysis was produced. These themes were regarded as IC-initiating moves. The same method was followed in order to obtain and categorize IC response strategies.

In order to determine the relationship between Iranian students' gender and the themes they tend to use as an IC, the relevant data were processed statistically with IBM SPSS Statistics 24. First, the data were tabulated through Crosstab Tables. Then, a Chi-square test of independence was carried out to determine whether the identified themes significantly differed between male and female students. In addition, Cramer's V test was performed to test the strength of the association.

Two significant notes are to be made here. First, as for the identification of the themes and the strategies, there was an inevitability concerning treating the data based on the literature and the researchers' own understanding of how IC initiating moves and their corresponding response strategies were realized in complaint occurrences in the university context among students. Nonetheless, an 'open category' was used, i.e., deliberate attempts were made to allow new patterns to emerge from the qualitative data. Doing so would obliterate the influence of the researchers' prejudgment about the data.



Second, as far as the reliability of the results was concerned, the data were analyzed by the researchers based on the definitions provided in *subsection 2.2.*, and Silverman's (2009) five approaches of refutational analysis, constant data comparison, comprehensive data use, inclusive of the deviant case and use of tables were employed to ensure the reliability of the results. They were then given to an experienced researcher in the field for analysis. In order to ensure maximum consistency between the researchers' analyses, Cohen's Kappa Coefficient was applied, the result of which indicated the degrees of agreement of 90.2% and 88.6% for the identified themes (IC initiating moves) and strategies (IC responses), respectively. All of the estimates were significant ( $p < 0.01$ ). Finally, the results of the analyses were put in juxtaposition, and the end result was reached by virtue of painstaking negotiations.

## 4. Results

### 4.1. Research question 1

The first research question was posed to identify the themes of Ics commonly used by Iranian students in the university context. As a result of analyzing the 239 IC exchanges of the data, 281 IC initiating moves were identified, which were categorized in light of three distinct themes, namely (1) *self*; (2) *other*; and (3) *situation*. Among these themes, *self*-Ics turned out to be the least frequent, making up only 9.6% of the total. *Other* Ics constituted 24.2% of the corpus, and *situation* Ics (Type A and Type B) were the most frequent, comprising as much as 66.2% of the total. Table 1 summarizes the frequency of each IC theme, as used by both male and female students, as well as their corresponding percentages:

**Table 1.** Descriptive statistics of the use of IC themes (IC initiating moves) by Iranian students (male and female) in the university context.

		Gender		Total	
		Male	Female		
IC Themes	Self	Count (n)	9	18	27
		% within theme	33.3%	66.7%	100.0%
		% of Total	3.2%	6.4%	9.6%
	Other	Count (n)	33	35	68
		% within theme	48.5%	51.5%	100.0%
		% of Total	11.7%	12.5%	24.2%
	Situation A	Count (n)	69	46	115
		% within theme	60.0%	40.0%	100.0%
		% of Total	24.6%	16.4%	40.9%
	Situation B	Count (n)	54	17	71
		% within theme	76.1%	23.9%	100.0%
		% of Total	19.2%	6.0%	25.3%
Total	Count (n)	165	116	281	
	% within theme	58.7%	41.3%	100.0%	
	% of Total	58.7%	41.3%	100.0%	

Ics that are used as an act of self-denigration are termed *self*-Ics. In comparison to other types of Ics, *self*-Ics were the least frequent, accounting for slightly less than  $\frac{1}{10}$  of the total corpus. In such Ics, both male and female students expressed negative evaluations about themselves; however, female students used twice as many *self* Ics as male students did. This clearly shows that these Iranian female students are more concerned about their appearance, personality, and physical and cognitive abilities than their male counterparts are. They are obviously not bashful about discussing their feelings about their own behavior, actions, and ability with their peers. As a matter of fact, they rejoice in *self*-Ics in which they denigrate themselves, for such Ics can be utilized to lessen social-relation discomfort degrees and, in turn, allow for commiseration and compliment. This is in contrast to the negative perspective of the face delineated by Brown and Levinson (1987). Interestingly enough, in this study, the most frequent type of strategy used by the students to respond to *self*-Ics (as will be discussed in the same section) was agreement or commiseration. However, it seems that the complaining speakers favored contradiction the most. All of this highlights the complex nature of the face and that it should be studied from different perspectives (Spencer-Oatey, 2007). An example of a *self*-IC follows:

- (a) *Self* IC. Two female students on the campus. (audio-recorded)

A: “*Ne-mi-dun-am čerâ inqadr zud asabâni mi-š-am.*”

NEG-PRS-know-1SG why so easily angry PRS-get-1SG

(I don't know why I get angry so easily.)

ب: به نظر من که خیلی هم آرامی.

B: “*Be nazar-e man ke xeyli ham ârum-i.*”

to opinion-ART I that very too calm-be.PRS.2SG

(In my opinion, you are very calm.)

In the above example, the speaker complained about her irritability. This IC was not left unanswered. Rather, her addressee tactfully contradicted her statement and built up a good rapport with the complaining speaker. In point of fact, only those instances of contradiction (the discussion of which will be provided later in the same section) that were provoked by *self*-Ics were appreciated by the speakers. Nevertheless, this was not far from the researchers' expectation since *self*-Ics generally focus on self-denigration and contradicting them functions as rapport-inspiring speech behavior.

The focus of *other* Ics is on other person/persons, especially on their personal or private affairs. This is synonymous with what is commonly known as 'gossip.' *Other* Ics made up slightly less than  $\frac{1}{4}$  of the total corpus. *Other* Ics are highly commonplace among students as they inspire rapport and confidence among conversationalists. As for this theme, the results, strangely enough, did not indicate any major difference between male and female students (Table 1). This struck the researchers, in particular, as odd that female students' *other* Ics (n=35) were not greater in number (at least not to a great extent) than those of their male counterparts (n=33). In fact, this finding might stand in contrast to the common belief that women tend to gossip more about other person/persons than men do, which is still commonplace in Iran.

The third identified theme of IC initiating moves is the *situation*, which is further divided into two subcategories: Type A *situations* and type B *situations*. Among other themes, *situation* Ics were the most frequent, constituting slightly more than 66% of the total corpus. Type A *situation* Ics are of **personal** focus; however, they are different from *self* Ics in that they are related to a specific situation at hand rather than a personal characteristic. On the other hand, the focus of Type B *situation* Ics is **impersonal**. It was only when the data were analyzed for the corresponding frequencies of the identified themes that the researchers began to notice that there was a considerable difference between males and females, whose uses of type B *situation* Ics made up 76.1% and 23.9% within its theme, respectively. One possible implication of this finding is that, when confronted with different options, these Iranian female students are less likely to participate in classroom discussions whose central focus is of global significance, such as political or economic issues. However, more studies are needed in this regard.

## 4.2. Research question 2

To find out the answer to the second question, the relevant data from Table 1 were analyzed by means of a Chi-square test of independence, which is also known as Pearson's Chi-square test. Table 2 displays the results of a 2x4 Chi-square analysis:

**Table 2.** Chi-square test of independence for the relationship between Iranian students' gender and the themes they tend to use as an IC.

	Value	df	Asymptotic Significance (2-sided)
Pearson Chi-Square	18.973 <sup>a</sup>	3	.000
Likelihood Ratio	19.426	3	.000
Linear-by-Linear Association	18.769	1	.000
N of Valid Cases	281		

a. 0 cells (0.0%) have expected count of less than 5. The minimum expected count is 11.15.

As Table 2 illustrates, the relationship between these variables is significant,  $\chi^2(3,281) = 18.973$ ,  $p < .05$ . In other words, the themes these Iranian students tend to use as an IC are not independent of their gender. In addition to the significance, it seemed useful to determine the strength of the association. Hence, Cramer's V test was performed as a correlation measure to test the strength of the association and, in turn, to discover how important the findings are. Table 3 presents the results of Cramer's V test:

**Table 3.** Results of Cramer's V test of association for the relationship between Iranian students' gender and the themes they tend to use as an IC.

		Value	Approximate Significance
Nominal by Nominal	Phi	.260	.000
	Cramer's V	.260	.000
N of Valid Cases		281	

The results of Cramer's V test ( $V = .26$ ) show that there is a moderate association between these Iranian students' gender and the themes they tend to use as an IC.

Furthermore, the observed frequencies of these Iranian students' preferred ICs (see Table 1) indicated that, comparatively, whereas female students tend to use *self*-ICs with a higher percentage (66.7%), male students tend to use *situation* ICs more often (66.1%) as opposed to their counterparts. Finally, as far as *other* ICs are concerned, the difference between males' and females' preferences was not very noticeable (48.5% and 51.5% for males and females, respectively).

### 4.3. Research question 3

The third research question aimed at exploring the major social strategies used by Iranian students to respond to ICs. As a result of a thorough thematic analysis, 321 IC responses with eight distinct types of strategies were identified among Iranian students at the university context while conversing with each other in Persian. It should be noted that two of these strategies emerged as new IC response strategies specific to the present context. The new IC response strategies that emerged in this study were coded as 'attribution' and 'admonishment,' comprising 11.8% and 10% of the corpus, respectively. Other strategies were coded as follows: (1)  $\emptyset$  response or topic switch; (2) questions; (3) contradiction; (4) joke/teasing (5) advice/lecture; and (6) agreement/commiseration. Among these strategies, the least frequent strategy was  $\emptyset$  a response or topic switch, making up 7.2% of the total.

In response to *self*-ICs, there was no  $\emptyset$  response or topic switch, and every IC, as an act of self-denigration, was responded to using one or some of the other IC response strategies. On the other hand, the most frequent strategy was agreement or commiseration, comprising 29% of the total. What is more, as a general trend, almost all of the IC response strategies were elicited by *situation* ICs with a higher percentage, with the only exception of the last category, i.e., 'admonishment,' which was elicited most often by *other* ICs in comparison with the other two types of ICs. Table 4 illustrates the exact percentages of each category:

**Table 4.** Descriptive statistics of the use of IC responses (IC response strategies) as a function of IC themes by Iranian students (male and female) in the university context.

		IC Themes			Total	
		Self	Other	Situation		
IC Response Strategies	$\emptyset$ or topic switch	Count (n)	0	4	19	23
		% within IC Responses	0.0%	17.4%	82.6%	100.0%
		% of Total	0.0%	1.2%	5.9%	7.2%
	Questions	Count (n)	2	10	22	34
		% within IC Responses	5.9%	29.4%	64.7%	100.0%
		% of Total	0.6%	3.1%	6.9%	10.6%
	Contradiction	Count (n)	10	3	18	31
		% within IC Responses	32.3%	9.7%	58.1%	100.0%
		% of Total	3.1%	0.9%	5.6%	9.7%
	Joke/teasing	Count (n)	5	13	24	42
		% within IC Responses	11.9%	31.0%	57.1%	100.0%
		% of Total	1.6%	4.0%	7.5%	13.1%
Advice/lecture	Count (n)	4	4	20	28	
	% within IC Responses	14.3%	14.3%	71.4%	100.0%	
	% of Total	1.2%	1.2%	6.2%	8.7%	
Commiseration	Count (n)	11	15	67	93	
	% within IC Responses	11.8%	16.1%	72.0%	100.0%	
	% of Total	3.4%	4.7%	20.9%	29.0%	

Attribution	Count (n)	9	11	18	38
	% within IC Responses	23.7%	28.9%	47.4%	100.0%
	% of Total	2.8%	3.4%	5.6%	11.8%
Admonishment	Count (n)	7	16	9	32
	% within IC Responses	21.9%	50.0%	28.1%	100.0%
	% of Total	2.2%	5.0%	2.8%	10.0%
Total	Count (n)	48	76	197	321
	% within IC Responses	15.0%	23.7%	61.4%	100.0%
	% of Total	15.0%	23.7%	61.4%	100.0%

Among the 321 IC response strategies, the least frequent one was  $\emptyset$  a response or topic switch. This means that the addressee either did not respond to the IC initiating move by the complaining speaker or tried to change the subject of conversation. Interestingly enough, in either case, the complaining speakers tried to talk their addressees into commiserating with them. By and large, it seems that in order to preserve rapport, these Iranian students tend to avoid topics with which they are prone to contradiction unless the complaining speaker brings up a direct follow-up question.

The second identified IC response strategy was questions. Unlike no response or topic switch, the questions strategy tempts the complaining speakers to elaborate on the complaint. Such questions are usually accompanied by nonverbal and verbal interjections such as ‘no way,’ or ‘really.’ It should also be noted that questions as an IC response strategy most often are not the ultimate response. Rather, they typically lead to other types of response strategies.

Another type of IC response strategy was coded as a contradiction, in which the addressee does not agree with the complaining speaker. This type of response, which serves as a disapproval strategy, in most cases tends to threaten the speaker’s face (except for *self*-ICs). Hence, the use of mitigating devices, such as hedges and indirect speech acts, among others, is not uncommon. Generally, it seems that when complaining about another person or situation, such Iranian students are inclined to maintain their solidarity with one another even if it means hiding their true feelings for the sake of appearances.

Joking or teasing refers to those IC responses that are intended to provoke laughter through funny and witty remarks. While instances of joking were not few in number, according to the observations by the researchers, rarely did they happen when strangers were present. Although this was not the aim of the present study (the relationship between social distance and the type of IC response strategies), it poses an intriguing question for further research.

Advice or lecture was another type of strategy identified in the corpus. It refers to those strategies by which the addressee gives his/her personal opinion or some suggestions about what someone had better do in a certain situation. Advice or lecture on how to behave or what to do in a particular situation by an addressee is highly appreciated by the complaining speaker since this strategy is typically perceived as rapport-inspiring speech behavior among friends. Besides, it shows that the addressee feels concerned about the comfort of the complaining speaker.

The most frequent IC response strategy was agreement or commiseration. In fact, the category of commiseration comprised 29% of the total corpus. Commiseration refers to any short exclamation or even long expression of feelings whose illocutionary force is to reassure the complaining speaker and make them stop worrying about something. One typical example of commiseration is when the recipient of the complaint, in sympathy with the complaining speaker, offers his/her condolences for a certain problem that has affected the complaining speaker. One interesting feature of such a strategy is that it usually serves as a conversational opener (Boxer, 1993b). The idea here is that only through explicit or embedded commiseration there is more potential for longer interactions among participants of conversations. In fact, sharing of common ideas and expressing attitudes based on the unfavorability of an offense is facilitated through such IC responses. Strictly speaking, the face of the participants, in this fashion, is almost never threatened but even enhanced. This happens through building up a good rapport while exchanging ideas. Moreover, the feeling of solidarity is strengthened among the conversationalists as everyone lets their guard down by divulging their actual stance on the occasion. This contrasts sharply with the traditional perspective on the face, particularly the FTAs (Brown & Levinson, 1987).

In addition to the abovementioned IC response strategies, two new strategies emerged in the present study, namely *attribution* and *admonishment*. Attribution in this study refers to a collection of responses whose common characteristic revolves around attributing the cause of a certain complaint to internal and/or external factors. Attribution also has to do with the philosophy or logic of a certain phenomenon, and they are usually accompanied by expressions such as “basically” and “technically.” To clarify the concept of attribution, let us look at an example:

(a) Three male friends in the corridor of the university building. (audio-recorded)

الف: کلاس بعد از نهار خیلی مزخرفه.

A: “*Clâss ba’d az nâhâr xeyli mozaxrafe-e.*”  
 class after lunch very crap-be.PRS.3SG  
 (After-lunch classes are crap.)

ب: آره آدم خوابش میگیره تو کلاس ...

B: “*Âre âdam xâb-eš mi-gir-e too clâss...*”  
 yeah human sleep-his PRS-get.PRS-3SG in class  
 (Yeah, it makes us drowsy in the classroom.)

ج: اصولا نباید سر ظهر کلاس باشه. بازدهی نداره اصلا.

C: “*Osulan na-bâyad sare zohr clâss bâš-e. Bâzdehi na-dâr-e aslan.*”  
 basically NEG-should at noon class be.PRS-3SG efficiency NEG-have.PRS-3SG at all  
 (Basically, there shouldn’t be any classes at noon. It’s not productive at all.)

In the above example, the speaker complains about the unfavorability of a certain situation. Two addressees respond to this *situation* IC using two different strategies. Whereas speaker B commiserates with the complaining speaker, speaker C attributes the unfavorable effect, ‘a boring class,’ to the relevant cause, ‘time of the day.’ Furthermore, he philosophizes about the consequence of this situation that “it’s not productive at all.” The linguistic realization of this strategy could even be formed using grammatical structures such as past modal. Speaker B, in example (c), responding with “You must have gone to bed late last night yet again.” is using the same strategy.

(b) Two female students are chit-chatting on the campus. (audio-recorded)

الف: من دیگه کلاس بعدی رو نمیتونم بیام.

A: “*Man dige clâss ba’di ro ne-mi-tun-am b-ia-m.*”  
 I anyway class next ART NEG-PRS-can-1SG PRS-come-1SG  
 (I can’t make it to the next class.)

خوابم گرفته.

“*Xâb-am gerefte.*”  
 sleep-my get.PTCP  
 (I feel sleepy.)

ب: حتما دیر خوابیدی باز دیشب.

B: “*Hatman dir xâb-id-i bâz di-šab.*”  
 definitely late sleep-PST-2SG again last-night  
 (You must have gone to bed late last night yet again.)

The other new strategy identified in the present study was coded as admonishment. This IC response strategy is used when the addressee finds the complaint of the complaining speaker insulting or patronizing, either to the self or other people. Admonishment means that the addressee disapproves of the complaining speaker's complaint. Following is an example of this IC response strategy:

(c) Two male students on the campus. (audio-recorded)

الف: دیدی چجوری راه میرفت؟ خود پنگوئن!

A: "Did-i      čejuri    râh      miraft?    Xode    panguan!"  
 see.PST-2SG    how      way      go.PST.3SG    itself      penguin  
 (Did you notice the way she walked? Exactly like a penguin!)

ب: زشته. مسخره نکن.

B: "Zešt-e.      Masxare    na-kon."  
 ugly-be.PRS.3SG    ridicule      NEG-do.IMP.2SG  
 (That's bad language. Don't mock people.)

In this example, the complaining speaker gossiped about another person (*another IC*). However, this was neither appreciated nor elaborated by the addressee. Instead, the recipient of the complaint admonished the complaining speaker for mocking people. Admonishment was the only category most frequently elicited by *other ICs* (50% within its total number). Therefore, it seems reasonable to assume that due to certain ideological backgrounds, part of which has been shaped by the Islamic context of Iran and the special norms of the university, some Iranian students are inclined to admonish 'ICs' in general, and '*other ICs*' in particular. This particular finding is in line with Goffman's (1967) call for due consideration concerning the particulars of the social context in which a given speech act takes place (Kalinina & Gabdreeva, 2020). In fact, this social dimension is to be added to our investigations of face-related speech acts.

## 5. Conclusion

The results of the study revealed that among the three identified themes, i.e., self, other, and situation, *self-ICs* were the least frequent and that female students used twice as many *self-ICs* as male students did. This indicates that these female students are more concerned about their appearance, personality, and physical and cognitive abilities than their male counterparts are. Moreover, a contradiction to *self-ICs* was highly appreciated by the complaining speaker, functioning as rapport-inspiring speech behavior. It was also discovered that, among other IC response strategies, only through explicit or embedded commiseration there is more potential for longer interactions among participants of conversations. This is, in part, related to the fact that commiseration typically serves as a conversational opener that leads to further sharing of ideas. What is more, as far as IC response strategies are concerned, two new response strategies were identified, namely attribution and admonishment. The conclusion to be drawn here is that in order to account for all of the strategies used by Iranian students to respond to ICs, the existing categorizations (see Boxer (1996) for such a categorization) are to be expanded.

The present study sheds some light on the pragmatics of the Persian language. In light of the present research, as well as the existing literature in the fields of speech act theory in general and complaints in particular, we can better understand the underlying mechanisms and components of ICs as used by Iranians in the Persian language, specifically in university contexts. Such findings are valuable since they aid us in understanding how ICs are realized in Iranian students' use of L1, the result of which provides new insights into the areas that Iranian students are more likely to make positive and/or negative transfer while learning a new language such as English. Similarly, such knowledge is vital for learners of Persian, "in that recognition of how native speakers use the speech act as well as knowledge of how to respond appropriately may open opportunities for non-native speakers to make friends" (Boxer & Pickering, 1995, p. 46). The significance of learners' mastery over pragmatic competence has been previously established (Zarei & Mohammadi, 2012). Additionally, this paper stands out among other similar studies,

due to the fact that the relevant data from students were not elicited but collected from spontaneous speech, naturally occurring in the university context. Finally, the major theoretical implication of this research is to support the notion that knowledge of speech acts and how to use them are noticeably culture-bound. As for the present study, the identified themes and strategies among students were, to a great extent, a reflection of the sociocultural norms of the Islamic context of Iran, as well as the special norms and standards of the university context.

This research had certain limitations with regard to sample size, data collection procedure, and instrumentation. The researchers are well aware of the threat that, despite the identified saturation, the results may not be generalizable to the whole university context. This is because the convenience sampling was performed on a small scale (only one university) due to feasibility considerations. Hence, further research could perform the same study on a larger scale. It should also be noted that the major instrument of the present study for data collection was through audio-recording. This procedure, albeit useful, could be triangulated with other instruments and data sources, such as role-plays (see, for example, Malmir & Taji, 2021), discourse completion tests (see, for example, Jalilifar & Hoseini, 2021), or interactions on social media platforms such as Twitter (see for example Depraetere et al., 2021), so as to enhance validity across data sources. Finally, as the present study was restricted to IC themes and IC response strategies, as well as the relationship between Iranian students' gender and their preferred IC themes in university contexts, future studies could investigate the effect of other factors, including social distance, on IC themes or replicate the same study in a different context, such as in the workplace.

## 6. References

- Allami, H. (2006). A sociopragmatic analysis of griping: The case of Iranian students. *The Linguistics Journal*, 1, 59-76.
- Archer, L., Hutchings, M., Ross, A., Leathwood, C., Gilchrist, R., & Phillips, D. (2003). *Higher education and social class: Issues of exclusion and inclusion*. RoutledgeFalmer.
- Arundale, R. B. (2006). Face as relational and interactional: A communication framework for research on face, facework, and politeness. *Journal of Politeness Research*, 2(2), 193-216. <https://doi.org/10.1515/PR.2006.011>
- Bedewy, D., & Gabriel, A. (2015). Examining perceptions of academic stress and its sources among university students: The perception of academic stress scale. *Health Psychology Open*, 2(2), 1-20. <https://doi.org/10.1177/2055102915596714>
- Boxer, D. (1993a). Complaints as positive strategies: What the learner needs to know. *TESOL Quarterly*, 27(2), 277-299. <https://doi.org/10.2307/3587147>
- Boxer, D. (1993b). Social distance and speech behavior: The case of indirect complaints. *Journal of Pragmatics*, 19, 103-125. [https://doi.org/10.1016/0378-2166\(93\)90084-3](https://doi.org/10.1016/0378-2166(93)90084-3)
- Boxer, D. (1996). Ethnographic interviewing as a research tool in speech act analysis: The case of complaints. In S. M. Gass, & J. Neu (Eds.), *Speech acts across cultures: Challenges to communication in a second language* (pp. 217-239). Mouton de Gruyter.
- Boxer, D., & Pickering, L. (1995). Problems in the presentation of speech acts in ELT materials: The case of complaints. *ELT Journal*, 49(1), 44-58. <https://doi.org/10.1093/elt/49.1.44>
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. <https://doi.org/10.1191/1478088706qp063oa>
- Brown, P., & Levinson, S. C. (1987). *Politeness: Some universals in language usage*. Cambridge University Press.
- Complaint. (n.d.). In *Merriam-Webster's collegiate dictionary*. Retrieved from <https://www.merriam-webster.com/dictionary/complaint>
- Decock, S., & Depraetere, I. (2018). (In) directness and complaints: A reassessment. *Journal of Pragmatics*, 132, 33-46. <https://doi.org/10.1016/j.pragma.2018.04.010>
- Depraetere, I., Decock, S., & Ruytenbeek, N. (2021). Linguistic (in) directness in twitter complaints: A contrastive analysis of railway complaint interactions. *Journal of Pragmatics*, 171, 215-233. <https://doi.org/10.1016/j.pragma.2020.09.026>

- Dersley, I., & Wootton, A. J. (2000). Complaint sequences within antagonistic arguments. *Research on Language and Social Interaction*, 33(4), 375–406. [https://doi.org/10.1207/S15327973RLSI3304\\_02](https://doi.org/10.1207/S15327973RLSI3304_02)
- Drew, P. (1998). Complaints about transgressions and misconduct. *Research on Language and Social Interaction*, 31(3-4), 295-325. <https://doi.org/10.1080/08351813.1998.9683595>
- Drew, P., & Holt, E. (1988). Complainable matters: The use of idiomatic expressions in making complaints. *Social Problems*, 35(4), 398-417. <https://doi.org/10.2307/800594>
- Drew, P., & Walker, T. (2009). Going too far: Complaining, escalating and disaffiliation. *Journal of Pragmatics*, 41, 2400–2414. <https://doi.org/10.1016/j.pragma.2008.09.046>
- Eslamirasekh, Z. (1993). A cross-cultural comparison of the requestive speech act realization patterns in Persian and American English. In L. F. Bouton, & Y. Kachru (Eds.), *Pragmatics and Language Learning* (Vol. 4, pp. 85-103). University of Illinois at Urbana-Champaign, Division of English as an International Language.
- Fogal, D., Harris, D. W., & Moss, M. (Eds.). (2018). *New work on speech acts*. Oxford University Press.
- Goffman, E. (1967). *Interaction ritual: Essays in face-to-face behavior*. Aldine Transaction.
- Goodwin, M. H. (1990). *He said–she said: Talk as social organization among black children*. Indiana University Press.
- Gundogan, S. (2023). The relationship of covid-19 student stress with school burnout, depression and subjective well-being: Adaptation of the covid-19 student stress scale into Turkish. *The Asia-Pacific Education Researcher*, 32(2), 165-176. <https://doi.org/10.1007/s40299-021-00641-2>
- Hathaway, E. D., Peyer, K. L., & Doyle, K. A. (2023). A first look at perceived stress in southeastern university students during the COVID-19 pandemic. *Journal of American College Health*, 71(2), 329-332. <https://doi.org/10.1080/07448481.2021.1895809>
- Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind. Revised and Expanded 3rd Edition*. McGraw-Hill.
- Holmes, J. (1990). Apologies in New Zealand English. *Language in Society*, 19(2), 155-199. <https://doi.org/10.1017/S0047404500014366>
- Holmes, J. (1995). *Women, men and politeness*. Longman.
- Holmes, J. (2006). *Gendered talk at work: Constructing gender identity through workplace discourse*. Blackwell.
- House, J., & Kádár, D. Z. (2023). A new critique of the binary first-and second-order distinction in politeness research. *Journal of Pragmatics*, 213, 145-158. <https://doi.org/10.1016/j.pragma.2023.06.001>
- Izadi, A. (2017). Culture-general and culture-specificity of face: Insights from argumentative talk in Iranian dissertation defenses. *Pragmatics and Society*, 8(2), 208-230. <https://doi.org/10.1075/ps.8.2.03iza>
- Jääskä, E., & Aaltonen, K. (2022). Teachers’ experiences of using game-based learning methods in project management higher education. *Project Leadership and Society*, 3, 100041. <https://doi.org/10.1016/j.plas.2022.100041>
- Jalilifar, A., & Hoseini, M. (2021). A cross-cultural study of compliment responses: A pragmatic analysis of a Persian and English speech act. *Language Related Research*, 12(5), 81-109. <https://dx.doi.org/10.52547/LRR.12.5.5>
- Jefferson, G. (1988). On the sequential organization of troubles talk in ordinary conversation. *Social Problems*, 35(4), 418-441. <https://doi.org/10.2307/800595>
- Kalinina, G. S., & Gabdreeva, N. V. (2020). The intention of “Apology” in speech etiquette: Based on the material of different structural languages. *Journal of Research in Applied Linguistics*, 11(Special Issue), 292-299. <https://doi.org/10.22055/RALS.2020.16323>





- Kozlova, I. (2004). Can you complain? Cross-cultural comparison of indirect complaints in Russian and American English. *Prospect*, 19(1), 84-105.
- Laforest, M. (2002). Scenes of family life: Complaining in everyday conversation. *Journal of Pragmatics*, 34(10-11), 1595-1620. [https://doi.org/10.1016/S0378-2166\(02\)00077-2](https://doi.org/10.1016/S0378-2166(02)00077-2)
- Leathwood, C., & O'Connell, P. (2003). 'It's a struggle': The construction of the 'new student' in higher education. *Journal of Education Policy*, 18(6), 597-615. <https://doi.org/10.1080/0268093032000145863>
- Leech, G. (2014). *The pragmatics of politeness*. Oxford University Press.
- Malmir, A., & Taji, N. (2021). The interplay of action, context, and linguistic vs. non-linguistic resources in L2 pragmatic performance: The case of requests and refusals. *Language Related Research*, 12(3), 215-253.
- Mills, S. (2003). *Gender and politeness*. Cambridge University Press.
- Murphy, B., & Neu, J. (1996). My grade's too low: The speech act set of complaining. In S. M. Gass, & J. Neu (Eds.), *Speech acts across cultures: Challenges to communication in a second language* (pp. 191-216). Mouton de Gruyter.
- Pomerantz, A. (1986). Extreme case formulations: A way of legitimizing claims. *Human Studies*, 9, 219-229.
- Rodriguez, A. (2022). "Ay no I do feel exhausted": Affiliative practices and interpersonal relationships in indirect complaints in Spanish. *Journal of Pragmatics*, 194, 39-53.
- Ruusuvuori, J., & Lindfors, P. (2009). Complaining about previous treatment in health care settings. *Journal of Pragmatics*, 41, 2415-2434.
- Schegloff, E. A. (1988). Goffman and the analysis of conversation. In P. Drew, & A. J. Wootton (Eds.), *Erving Goffman: Exploring the interaction order* (pp. 89-135). Cambridge University Press.
- Schegloff, E. A. (2005). On complainability. *Social Problems*, 52(4), 449-476.
- Silverman, D. (2009). *Doing qualitative research* (3rd ed.). Sage.
- Spencer-Oatey, H. (2007). Theories of identity and the analysis of face. *Journal of pragmatics*, 39(4), 639-656.
- Tatsuki, D. H. (2000). If my complaints could passions move: An interlanguage study of aggression. *Journal of Pragmatics*, 32, 1003-1017.
- Traverso, V. (2009). The dilemmas of third-party complaints in conversation between friends. *Journal of Pragmatics*, 41, 2385-2399.
- Tseng, M. Y., & Chen, H. (2022). Mutual face-maintaining acts: An analysis of talks between NBA referees and players/coaches. *Journal of Research in Applied Linguistics*, 13(2), 44-52.
- Ulriksen, L., Madsen, L. M., & Holmegaard, H. T. (2017). The first-year experience of non-traditional students in Danish science and engineering university programmes. *16*(1), 45-61.
- Zarei, G. R., & Mohammadi, M. (2012). E-politeness in Iranian English electronic requests to the faculty. *Journal of Research in Applied Linguistics*, 3(1), 3-24.



<https://doi.org/10.22077/ali.2023.6799.1020>

## Reliability and Factorial Study of Writing Self-Regulation Inventory in Iranian EFL Context

Husain Abdulhay<sup>1\*</sup>   
 Moussa Ahmadian<sup>2</sup> 

<sup>1</sup>Lecturer in Applied Linguistics, Department of English Language, Faculty of Literature and Foreign Languages, Arak University, Arak, Iran

<sup>2</sup>Associate Professor in Applied Linguistics, Department of English Language, Faculty of Literature and Foreign Languages, Arak University, Arak, Iran

### ABSTRACT

Self-regulated learning is leading the way as a self-sustaining technique helping learners go through with a task and make headway. This research attempted to examine the underlying construct of writing self-regulation inventory extrapolated to Iranian EFL university context. A total sample of 116 sophomore EFL university learners attending an essay writing course participated in this study. They were asked to do an in-class writing task assignment and immediately fill in an inventory. Pearson product-moment correlation and confirmatory factor analysis procedure using principal component analysis were applied to ascertain the relationships among the variables and construct validity of the instrument, respectively. The inventory yielded a strong internal consistency as to make the results of the present study reliable and to prove replication studies justifiable. The highest overall means and internal consistency reliability estimates were recorded both for the goal subscale. However, the achievement on task was negatively associated with cognitive, meta-cognitive, motivational, and behavioral dimensions of writing self-regulation. The results of factor analysis disclosed that goal, efficacy, meta-cognitive, cognitive, motivational, and behavioral variables are loaded on a single component; that is, they measured the same underlying theoretical construct. On the whole, the findings imply that the construct of writing self-regulation is consistent in EFL university context.

**KEYWORDS:** Academic writing; EFL learners; Reliability; Writing self-regulation

### ARTICLE HISTORY

Received: 10 April 2023  
 Revised: 16 June 2023  
 Accepted: 22 July 2023  
 Published: 30 September 2023

### CORRESPONDING AUTHOR

E-mail: [husainabdolhay@yahoo.com](mailto:husainabdolhay@yahoo.com)

## 1. Introduction

Self-regulated learning has turned over a new leaf by assisting learners to untangle turns and twists in a course of action toward goal attainment. Self-regulated learning, in its broad concept, is expounded as an iterative, self-steering, and multi-layered process that targets one's cognition, action, and affection, as well as features of the environment for modulation in the pay of one's own goals (Boekaerts et al., 2005). Bernacki et al., (2014) know of self-regulation as meta-cognitive tasks, requiring adequate motivation to initiate and maintain engagement. As Winne et al., (2002) postulate, self-regulation is a situation-related act; that is, the task at hand exposes learners to self-regulation. It enables to handle learning to the utmost of ability to realize goals irrespective of barriers and blockades, either inner or outer issues.

Writing is a complex process that conveys the writer's knowledge, skill, and tactics, which seeks motivation and self-regulation (MacArthur et al., 2015). Writing processes are conspicuously, in essence, self-regulatory, viz., planning textually, establishing goals, revamping, organizing, and redacting, and revamping (Hidi & Boscolo, 2006). Zimmerman and Kitsantas (2007) contend that successful writing seeks a great extent of self-regulation and self-motivation. Ruan (2005) notes that the research on writing, within the cognitive framework, shows that, the skilled writers' composition processes embrace extensive self-regulation and meta-cognitive control.

As Flower and Hayes (1981) put it, success in writing is pertained to the extent that writers have the ability to regulate their writing process. This relation explains the reasons why many instructors focus on the promotion of self-regulatory strategies like planning, reviewing, and editing for advancement of the composition process (Harris & Graham, 1996). Writing self-regulation is the ability to write effectively, using writing instructional tools without regular direction and/or prompting (Schunk & Zimmerman, 2007). Zimmerman and Kitsantas (2007) contend that successful writing seeks a great deal of self-regulation and self-motivation.

Writing self-regulation is theorized under a triad of approaches: cognitive, socio-cognitive, and socio-cultural perspectives. Cognitive perspective was explained by Flower and Hayes (1981) and Hayes (2012), as a problem-solving act comprising a set of complicated planning, translation, and evaluation processes, with regulation being a key factor to monitor one's own cognitive writing behavior. Socio-cognitive perspective suggested by Zimmerman and Schunk (1989) supplements the former cognitive models, the emotion, context, and personal behaviors like motivation, task objectives, and instructional feedback through which writers communicate with each other to modify their output and accommodate their writing self-efficacy beliefs (Sala-Bubare & Castello, 2018). The socio-cultural perspective, reversely, accentuates the importance of the institutional output and socially and historically situated cognition; that is, the socially participated and mediated processes (Prior, 2006), by which meaning is built, modified and transformed, and co-regulated, with the aid of a more skilled writer (Allal, 2018). However, Hidi and Boscolo (2006) point out that motivational constructs in writing can be subdivided into three main domains: self-belief (self-efficacy and self-concept); motives to perform (interest and perceived value in writing), and writing self-regulation.

Goal plays as a road map availing person of pathways as to enter to reach his/her end point. Kaplan et al., (2009) point out that most investigations describe achievement goals and self-regulation as distinct, albeit related, constructs. Performance-goal-orientated learners take unsuccessful attempts as reflections of their ability and would consequently discard effort to learn the new material (Dweck et al., 1989). Students with performance goal orientation are concerned about how others judge their ability. They look after others' approvals (Canfield & Zastavker, 2010). They are concerned about social comparison (Cheung, 2008), seeing people as reference group to outperform in a task. Such students cannot consider themselves successful unless they compare themselves with others (Dinc, 2010). Performance-avoidance-oriented students seek to escape situations that make them appear deficient in ability or less capable than others (Wolters, 2004).

Ruan (2005) notes that the research on writing within the cognitive framework evidences that the skilled writers' composition processes embrace extensive self-regulation and meta-cognitive control. Writing self-regulation is also related to students' self-efficacy to perform writing tasks (Zumbrunn et al., 2016). However, cross-cultural tapping and model replicating of self-regulated learning for the internal consistency and reliability underlying its component constructs in the context of EFL classroom structures is sought. Contingent on the aforesaid points and the upshots of the studies being reviewed above, this study endeavored to bridge the void as per the construct of writing self-regulation between Iranian EFL and target-language contexts. Thus, to assay the factor structure and psychometric properties of writing self-regulation inventory among Iranian EFL learners, the following question was constructed to be answered:

*Q:* Does factorial structure of writing self-regulation demonstrate evidence of internal consistency and reliability in Iranian EFL context?

Likewise, pursuant to the posed research question, the following hypothesis was formulated:

*H:* The factorial structure of writing self-regulation empirically demonstrates substantial evidence of internal consistency and reliability in Iranian EFL context.

## 2. Literature Review

Self-regulation is a self-sustaining technique serving individuals to mend their ways and make headway. A burgeoning body of research prevails that deals with self-regulated learning. Malpass et al., (1999), through structural equation modeling techniques, examining the relationship between math achievements of high school students and their self-regulation, self-efficacy, goal orientation, and worry found that learning goal orientation, commonly called as mastery goal orientation, was positively related to students' self-regulatory behaviors. Meece and Miller (1999), in a study, exploring writing goal orientations of 431 elementary learners, evidenced that goal orientations and the use of writing self-regulation strategies were positively correlated to writing mastery goals and negatively associated with writing performance-avoidance goals. Pajares and Cheong (2003), examining the achievement goals and writing self-efficacy of 1266 K-12 (kindergarten to 12th grade)

school students (aged 9-17), evidenced positive relationships between mastery goals and students' self-efficacy for self-regulation.

Linnenbrink (2005), in a quasi-experimental study, during five weeks, examined the impacts of goal-conditioned math classroom (mastery, performance-approach, spliced mastery/performance-approach goals) and personal goal orientations on 237 upper elementary students' motivation, emotional well-being, help-seeking, cognitive engagement, and achievement. The classroom goal condition had a significant effect on help-seeking and achievement, while the spliced condition proving the most beneficial pattern. Personal mastery goals were innocuous for eleven of twelve upshots including achievement; personal performance-approach goals were nocuous for achievement and test anxiety and disconnected to the rest of outcomes. Based on the entering personal goal orientation, the effect of the classroom goal condition did not change. Mastery goals benefited the learners while performance-approach goals were debilitating to the achievement.

Kozlowski and Bell (2006), investigating a large sample of college students' goal orientation in relation to self-regulation, found that mastery and performance goals had a link with cognitive self-regulatory activity, whereas learning goals are significantly linked with upper self-regulatory practice. Kitsantas et al., (2009) conducted a study on 81 fifth graders' academic achievement contingent upon their erstwhile achievement, self-regulation strategy use, and goal. The findings indicated significant variance of prior achievement and self-regulation predicting achievement, whereas performance upshots across diverse courses are not significantly accounted for by goal orientation. Self-regulation is endorsed to fall utile in context of second language learning for fertile learning (Harrison & Prain 2009). However, Tseng et al., (2006) substantiated applicable extrapolation of self-regulation to vocabulary domain at university and high school levels.

Kozlowski and Bell (2006), investigating the relationships between goal orientation and self-regulation of a large sample of college students, observed that mastery and performance goals were positively related to cognitive self-regulatory activity, whereas learning goals were significantly associated with higher self-regulatory activity. Harris et al., (2008) regarded self-regulation as a sine qua non for a writer to appear goal-oriented, reflective, and resourceful by bringing into work cognitive processes and appropriate strategies for planning, text production, and revision.

Pratontep and Chinwonno (2008) scoured 30 Thai university learners' reading self-regulatory tactics. Frequent use of meta-cognitive and performance regulation tactics were evidenced for group of lower level while group with higher level illustrated active uses of self-regulated learning tactics for adjustment of their meta-cognition and performance more often than the lower level did. Additionally, verbal protocols evinced more self-regulated tactics in performance and volitional than forethought or self-reflection phases.

Al-Harthy et al. (2010) probed relation of students' total scores on 12 exams for academic achievement to self-efficacy, task value, goal orientations, meta-cognitive self-regulation, and self-regulatory learning strategies. Results of path analysis proved achievement mastery goals directly and positively influence deep learning, meta-cognitive self-regulation, and self-regulatory strategies. Likewise, a strong relation of task-value, self-efficacy, self-regulation, and elaboration to academic achievement was recorded. Academic achievement directly was not significantly affected by mastery goals, meta-cognitive self-regulation, and deep learning strategies. Scores strongly, mastery goals positively, but avoidance goals were negatively accounted for by self-efficacy.

Bernacki et al., (2012), in a structural equation modeling analysis, explored goal orientations and self-regulation behaviors of college students exposed to technology designed to enhance reading comprehension. The results indicated that mastery goal students had more proclivity for information attainment, note-taking, and acquisition monitoring. Performance-approach goals did not predict self-regulation behaviors, and performance-avoidance orientation was a negative predictor of note-taking and information-seeking.

El-Henawy et al., (2012) explored the effectiveness of the self-regulated strategy development instruction on thirty junior teachers' argumentative writing. The upshots disclosed the quality of argumentative essays in three domains viz., content, organization, and argumentation improved in all experimental groups. Development of the students' writing performance was preserved over time and generalized to untaught writing genres. Syafitri (2020) evidenced that senior eleventh-grade high-school students' self-regulation and writing skills were not correlated.

### 3. Method

A cross-sectional method was adopted for conducting the present study in the form of immediate administration of self-report inventory after taking a sample of the university students' in-class essay performances. To make sure the learners had undergone enough of argumentative essay-writing instructions, the one-shot single-session data collection from different classrooms with different teachers was scheduled for the final weeks of the spring semester. Ensuring the confidentiality of their responses, the questionnaire administration came immediately after they had completed their five-paragraph argumentative essay task during their regular class time. The data gathered out of this cross-sectional study were

imported to SPSS 20 software for examining the reliability estimates and principal component analysis to apprehend the underlying factor structure of the inventory.

### 3.1. Participants

The data were collected from a sample of 116 sophomore undergraduate state-run university learners, aged 19-23 years (37/6% male & 62/4% female) attending essay writing course. The sampling was based on the accessibility of classrooms for single-session assignment of an essay writing task followed by questionnaire administration. The participants had already undertaken three two-credit courses, just to name them, English Grammar (1) and (2), and Paragraph Writing. Notably, the Oxford Quick Placement Test (OQPT) (2004) was administered to make sure of the homogeneity of the learners in the light of general language proficiency so as to reduce variations of data for the intended analyses. As for English language proficiency, intermediate level the attendees were ranked based on their scores one standard deviation above and below the mean.

### 3.2. Instruments

A self-report instrument developed by Lichtinger et al., (2006), anchored by 1 “not at all true of me” to 7 “very true of me”, was administered in this study (See the Appendix). This measure was designed by validating through correlation with Motivated Strategies for Learning Questionnaire originated by Pintrich et al., (1991), to rate various employments of cognitive, meta-cognitive, motivational, and behavioral strategies in writing. It also comprised efficacy and goal orientation subscales adapted from Patterns of Adaptive Learning Survey developed by Midgley et al. (2000).

Meta-cognitive component, which assesses the use of strategies that help students control and regulate their own cognition, consists of 7 subsections. The motivation section includes three measures: task value encouragement, success encouragement, and administering self-praise. Behavioral segment includes only help-seeking subscale. The cognitive dimension includes reader awareness, eliciting context, and verbalization strategies. The efficacy subscale includes sample items. The goal orientation dimension aims to measure seven subcomponents.

### 3.3. Data collection

Data collection was based on the administration of a self-report instrument immediately following the participants' completion of an in-class essay task. The instructions on how to approach the writing performance were given by one of the researchers. The students were assigned to write within 40 minutes an argumentative essay on “Gender Differences” in the classroom. They were directed to write at least a five-paragraph argumentative essay (including a general introduction paragraph, three detailed body paragraphs, and a general concluding paragraph). Likewise, they were offered to consider issues like sports, social life, and co-education in their arguments. To promote their participation, they were assured of the confidentiality of their information. The sampling was completed within one-session class time under the supervision of one of the researchers who gave guidance and assistance.

### 3.4. Scoring of writing

International English Language Testing System (IELTS) rubric (Shaw & Falvey, 2008) was adopted for guiding the rating of the overall quality of the essays in this study. This scoring profile is an extensively well-used and -researched benchmark for the assessment of an individual's ability in writing of clear formal English as the requirement in academic settings (Uysal, 2009). Through the application of an internationally recognized standardized grading protocol like the IELTS, this study sought to increase the external validity. Two raters, both with PhD degree in Applied Linguistics, were invited to rate the essay performances. The average of the scores obtained from the raters was considered as the final score for each participant for data analysis.

## 4. Results

Cronbach's alpha reliability was utilized to identify the internal consistency of the subcomponents of writing self-regulation inventory. The reliabilities of all the variables are presented in Table 1.

**Table 1.** Internal Consistency Reliability of the Factors

Cronbach's Alpha	Number of Items
------------------	-----------------

Goal	.917	34
Efficacy	.837	5
Metacognition	.847	23
Motivation	.854	8
Cognition	.636	9
Behavior	.865	4

Cronbach's alpha coefficients were analyzed to assess the internal consistency reliability of each subscale. The reliability analysis using Cronbach's alpha coefficients revealed that the overall internal consistency reliabilities of all the subscales were substantially adequate. Reliability analysis indicates that all alpha coefficients were well within the acceptable range, with actual values of .63 to .91, thereby verifying the reliability of the constructs. As indicated in Table 1, the results exhibit the strongest overall internal consistency for goal subscale (.917). The results of Cronbach reliability estimates exhibit a strong overall internal consistency for the meta-cognitive, motivational, behavioral (.847, .854, & .865, respectively), and cognitive ( $\alpha = 0.636$ ) subcomponent variables. Descriptive statistics, including mean and standard deviations of all the subcomponents, are reported in Table 2.

**Table 2.** Descriptive Statistics of the Scale Variables

	Means	Std. Deviation
Goal	154.2537	34.38016
Efficacy	25.8103	6.51141
Metacognitive strategies	114.7586	23.38696
Motivational strategies	40.1724	9.75908
Cognitive strategies	37.6034	10.40642
Behavioral strategies	18.7586	6.86468

Based on the results of descriptive statistics, the mean score of goal variable was the highest (154.2537). However, the mean score of meta-cognition (114.7586) was much higher than that of motivation, cognition, and behavior categories (40.1724, 37.6034, and 18.75.86, respectively). This suggests that the students were more meta-cognitively involved. Likewise, the standard deviation of meta-cognition scores was higher than that of the rest, indicating that meta-cognition scores were more widely spread.

Confirmatory factor analysis using principal component analysis procedure was carried out to examine the structural construct validity of model of writing self-regulation composed of goal, efficacy, and cognitive, meta-cognitive, motivational, and behavioral variables. Before doing this, Kaiser normalization and Bartlett test of Sphericity were calculated so as to identify the data's appropriateness for the factor analysis. With respect to the data's adequacy for the factor analysis, the Kaiser-Meyer-Olkin (KMO) coefficient index needs to be greater than .50 and the Bartlett Sphericity test has to be significant (Field, 2013; Kaiser, 1974). The factorial validity results for the items of the measurement tool are reported below. Table 3 displays the results of the Kolmogorov-Smirnov and Bartlett sphericity tests.

**Table 3.** KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy	.811
Bartlett's Test of Sphericity Approx. Chi-Square	248.913
df	15
Sig.	.000

As shown in Table 3, the value of the Kaiser-Meyer-Olkin Measure of Sampling Adequacy was .81, demonstrating a high level of inter-correlations among the factors. Furthermore, KMO index (0.811) and also the Bartlett sphericity test ( $X^2 = 248.913$ ,  $df = 15$ ,  $p < .000$ ) showed to be significant, providing evidence that the data collected were appropriate for performance of factor analysis. Table 4 lists the upshots.

**Table 4.** Squared Loadings Extraction Sums

Component	Eigenvalues			Squared Loadings Extraction Sums		
	Total	Variance	Cumulative	Total	Variance	Cumulative
1	3.163	52.716	52.716	3.163	52.716	52.716
2	.916	15.264	67.979			
3	.653	10.892	78.871			
4	.597	9.943	88.813			
5	.368	6.131	94.944			
6	.303	5.056	100.00			

Table 4 shows that one component was only extracted with eigenvalues 3.163, meeting the criterion of being larger than 1 as for a factor to be acceptable, and accounting for 52.716 percent of the total variables, which is sufficient as an acceptable variance ratio needs to be between 40 and 60 percent (Tavşancıl, 2014).

The factorial structure and construct validity of the inventory subcomponents was explored through principal component analysis using varimax rotation in which the minimum-eigenvalue criterion of 1.00 was set as the threshold to determine the number of factors and find underlying pattern of relationships between the intended variables. Table 5 evinces the upshots.

**Table 5.** Component Matrixa

	Component
	1
Goal	.739
Metacognition	.817
Motivation	.857
Cognition	.679
Behavior	.589
Efficacy	.638

*Extraction Method: Principal Component Analysis.*

*a. 1 component extracted.*

*Rotated Component Matrix<sup>a</sup>*

*a. Only one component was extracted. The solution cannot be rotated.*

Based on Table 5, the variables display loadings on one component. Principal component analysis with varimax rotation ended up in a one-factor solution for the goal, efficacy, cognitive, meta-cognitive, motivational, and behavioral variables. All of these subcategories were loaded on a single factor, accounting for 52.716 percent of the total variance. The fact that 52.716 percent of the scale was explained is deemed to be sufficient since an acceptable variance ratio should come between 40% and 60% (Tavşancıl, 2014).

Confirmatory factor analyses using principal component analysis with varimax rotation indicated a six-factor final solution. On a practical note, Cronbach's alpha value for goal subscale was higher than the other ones. Cronbach's alpha estimate for help-seeking variable was ranked second in order of magnitude. However, Cronbach's alpha reliability coefficient for cognitive subcomponent was lower than the rest. The results of principal component analysis established the hexa-factor model exhibiting the best fit to the data. PCA results portrayed that all the six sub-factors examined had moderate to high factor loading values, ranging from 0.589 to 0.857, to the corresponding factor. The outcomes proved that the instrument makes one component construct forming six factorial sub-dimensions.

Table 6 illustrates the results of descriptive analyses of goal, writing self-regulation, and task achievement.

**Table 6.** Descriptive Statistics of Self-regulation, Goal, &Task Achievement

	Task Achievement	Self-regulation	Goal
Mean	6.1963	213.8507	154.2537
Std. Deviation	1.46971	39.23680	34.38016

Variances	2.160	1539.526	1181.995
-----------	-------	----------	----------

According to the results of descriptive statistics, the mean score of total self-regulation (213.8507) was much higher than that of both goal orientations (154.2537) and grade (6.1963). This suggests that the students' concerns for self-regulation were higher than goal orientations. Likewise, the standard deviation of self-regulation scores was higher than that of the rest, indicating that self-regulation scores were more widely spread.

In order to discern the strength and direction of the association between the components of the inventory examined in the present study, component correlation was computed. Table 7 showcases the results.

**Table 7.** Correlation Matrix

Goal	-						
Metacognition	.481**	-					
Motivation	.623**	.652**	-				
Cognition	.422**	.394**	.470**	-			
Behavior	.330**	.401**	.387**	.407**	-		
Efficacy	.298**	.548**	.490**	.303**	.148	-	
Grade	.048	-.070	-.007	-.008	-.092	.067	-

Note. \*\* At the .01 level correlation is significant (2-tailed).

Table 7 showcases the associations among the variables. Accordingly, the strongest association can be observed between meta-cognition and motivation (652,  $p < .01$ ). A strong correlation can be also seen between motivation and total goals (.623,  $p < .01$ ). Goal was not significantly correlated with task achievement (.048). Efficacy was correlated with meta-cognitive, motivational, and cognitive subcomponents. However, efficacy was not significantly related to both help-seeking and task achievement (.148 and .067, respectively). Help-seeking, except for efficacy, was correlated with the other variables of self-regulation in writing. Goal, cognitive, meta-cognitive, and motivational variables were significantly correlated with each other. In the light of the findings from the correlation analysis, a negative correlation was evidenced between meta-cognitive, motivational, cognitive, and behavioral ( $r = -.070$ ,  $r = -.007$ ,  $r = -.008$ , and  $r = -.092$ ) and achievement on task. The greater the report of self-regulation in writing, the lower the achievement on task of academic writing in foreign language is, whereas the size of the value of the correlation coefficient was found not to be statistically significant.

## 5. Discussion

The purpose of this study was to measure the reliability and psychometric properties of writing self-regulation inventory embracing goal, efficacy, and cognitive, meta-cognitive, behavioral, and motivational dimensions. The focus of the study design was on the results of Cronbach's alpha and the principal component analysis (PCA). Also, the study aimed to assay the construct validity of the six-factor structure of writing self-regulation. The results denoted that the proposed hypothesis for this study can be supported. The theoretical model accounting for the self-regulation in writing was found to be a good fit to the observed data. The outcomes demonstrate that the instrument forms a component construct listing six subscales with high internal consistency reliability in Iranian academic writing context. Thus, the application of the instrument in this study yields support for the theoretical foundation for the cross-cultural function of self-regulation across writing domain.

The overall internal consistency of the inventory was found to be satisfactory, recording a high value for the observed variables. These findings take note of the fact that self-regulation in writing construct is multi-layered, embracing several independent variables. The results indicated that all six dimensions constituting the instrument were distributed into one component construct for the purpose of appraisal of Iranian EFL learners' writing self-regulation features. The psychometric quality of the subscales showed that this instrument is cross-culturally valid and reliable for the determination of the level of writing self-regulation practice of the university learners in Iranian context, which literally warrants the precision of upcoming research in EFL writing settings.

As regards the results of the validity study, the factors tied to the theoretical construct of writing self-regulation were found to be well-knotted into one component. The findings parallel to the study by Kanlapan and Velasco (2009), contextualizing and validating 115-item scale of self-regulation in written communication skills, endorse the adapted model of self-regulated learning (Pintrich, et al., 1993), extended to the domain-specific context of Iranian EFL writing. A one-factor structure clustering all the variables into one component construct made a good fit for the observed data. To put it another way, the findings indicated that the intended measurement variables were all unified making common cause in



gauging the same construct. The factor structure of the writing self-regulation served adequately to appraise the level of writing self-regulation of Iranian EFL learners, congruent with the study by Lichtinger et al., (2006), making it plausible to proffer strong empirically-based evidence on the applicability of this construct to Iranian EFL writing context. Taken together, the results confirm that the writing self-regulation extrapolated to Iranian EFL context elicit similar pattern in course of writing self-regulation analogous to the original measurement.

The results of factor analysis disclosed that goal, efficacy, meta-cognitive, cognitive, motivational, and behavioral variables are loaded on a single factor; that is, they, together, multi-dimensionally tested the same underlying construct. Moreover, the results provided cultural and contextual support for the stability and applicability of one-factor model of writing self-regulation; that is, goal, efficacy, meta-cognitive, cognitive, motivational, and behavioral variables are collapsed into one unitary construct. The self-regulatory factors were found to be closely threaded with goal ordination on which development and sustainability of writing behaviors are based. The strong loading of goal orientation listed in writing self-regulation framework does yield support for the theoretical perspective of achievement goal as an inalienable ingredient of self-regulation practice. The findings from factor analysis convey that teachers seeking to uplift and revamp learners' motivated writing self-regulation need to act out to boost their planned and goal-related behaviors as to their engagements in the subject matter.

It is encouraging to note that the highest overall mean score was observed for the goal subscale and the lowest one was evidenced for the help-seeking as a behavioral strategy. The Iranian EFL learners reported frequent goal endorsements for self-regulation of their writing. It sounds that goal-orientation serves as feeding source for learners to adjust their motivations, cognitions, meta-cognition, and behaviors to successfully fulfill the writing task. Based on the results of descriptive statistics, the mean score of meta-cognition (114.7586) was much higher than that of motivation, cognition, and behavior (40.1724, 37.6034, 18.75.86, respectively), consistent with Bailey's (2016) work evidencing the association between meta-cognitive strategies for writing and self-regulated learning behavior. This suggests that the students were more meta-cognitively involved. Likewise, the standard deviation of meta-cognition scores was higher than that of the rest, indicating that meta-cognition scores were more widely spread.

Significant correlational findings were released from the current study. The outcomes of correlation analysis demonstrated positive relationship among all the subscales. The findings provided ecologically valid data on writing performances in academic settings in support of a general model of self-regulation. The findings bolstered the theoretical foundations for self-regulation that acknowledge the integration of goal, self-efficacy, cognitive, meta-cognitive, motivational, and behavioral variables in its construct.

Iranian EFL learners in this study demonstrated to be fairly cognitively-engaged in trying to figure out their readers envisage and visualize things about which they wrote, and speak out the words they put into print through the use of reader awareness, eliciting context, and verbalization, respectively. Magno (2008) evidenced cognition-regulation is a significant predictor of written proficiency. The findings for the cognitive variable provide ecologically valid empirical data on academic performance on actual classroom in support of a general model of self-regulation in writing. Cognitive strategies can be scaled as the yardstick of proficiency in writing. Self-regulation in writing inventory can be exploited to palpate learners' performances for the use of cognitive strategies to prescribe remedial teaching.

The Iranian EFL learners' scores on the uses of meta-cognitive strategies were related to their goal orientations, supporting the study carried out by Al-Harthy and Was (2010) who evidenced the direct positive effect of achievement mastery goal on meta-cognitive self-regulation in a path analysis study. The highest correlation of relationship was reported between meta-cognition and motivation variables. This outcome suggests that the higher the level of motivation the learners enjoy, the more they are expected to exercise more control over their cognitive processes.

Based on the results of Pearson correlations, motivation was demonstrated to be significantly linked with all of the subcomponents of goal orientations and writing self-regulation. This suggests that students enthusiastically knuckle down to cognitive, meta-cognitive, and behavioral strategies, corresponding with the study by Teng and Zhang (2016) reporting that motivational regulation strategies such as interest enhancement, performance self-talk, mastery self-talk, emotional control, and environment structuring are antecedents of the perceived use of self-regulation tactics, particularly cognitive and meta-cognitive strategies, in EFL students' writing performances. The strong link between the motivation and goal orientation scales reminds teachers of the significance of the goal-situated setting in contribution to learners' motivations and as a result self-regulation of their writing.

The results of the study showed that efficacy significantly had the highest association with meta-cognition. It appears that students who believed they were capable of learning are more likely to report the use of cognitive and meta-cognitive strategies. These findings are in line with the studies by Pintrich and Garcia (1991) and Zimmerman and Bandura (1994), evidencing that students with high efficacy beliefs apply more cognitive and meta-cognitive strategies. This relation also corroborates the findings of the study by Bouffard-Bouchard et al., (1991) showing that self-efficacy significantly and positively influences the occurrence of some aspects of self-regulation like monitoring of working time and resolving conceptual problems. The results of the present study also support the study done by Artino and Stephens (2007) in online

courses that self-efficacy is significantly related to students' use of elaboration, critical thinking, and meta-cognitive learning strategies. Self-efficacy in this study was found to relate to goal orientation, corresponding to the study done by Pajares et al., (2000), which evidenced that writing self-efficacy is associated with learners' achievement goal orientations.

This study revealed some significant findings in terms of the relationship between Iranian EFL learners' goal orientations and writing self-regulation. The significant relationship observed between self-regulation in writing and goal orientation implies that writing self-regulation and goals are intermingled to control and steer students' writing performances. As Elliot (1999) proposes, goal orientation represents cognitive-dynamic forms of self-regulation. When students are goal-oriented to perform writing tasks, they are more likely to adjust their motivations, cognitions, meta-cognition, and behaviors to successfully fulfill the writing task. It can be inferred that the goals EFL learners set have a directive effect on their practices of writing self-regulation. Thus, Iranian EFL learners' writing self-regulation proves to be closely tied to their goal orientation. This suggests that students' levels of writing self-regulation are closely geared to the goals toward which they endorse. The consistent link between the goal orientation scale and all the other five self-regulation scales suggests that teachers can benefit from building a goal-conditioned environment to encourage students' stays on the subject matter. Taken together, the findings of the present study are also in line with the statement made by Elliot (1999) that goals represent cognitive-dynamic forms of self-regulation.

The findings also show that students' goal orientation is associated with their self-regulation in writing in the academic setting, which is in line with the results of previous studies (Kaplan et al., 2009; Farsani et al., 2014). The outcomes re-echo the clangor of confluence of goal-setting and self-regulation in advancement of writing practice. The outcomes highlight the integral value of multiple theoretical frameworks functioning in the cause of self-regulation; that is, socio-cognitive and achievement goal perspectives, represented in the model of self-regulation and goal orientation, respectively.

These findings provide insight regarding the importance of learners' goal endorsement in propulsion of motivational, cognitive, meta-cognitive, and behavioral strategies. The findings especially endorse the suitability of the intended goal types to be situated in the process of self-regulation in writing. Moreover, these findings imply that any self-regulation attempt is doomed to come up with failure unless a least degree of goal endorsement comes into effect. No matter how writing activities are hampered by obstacles, interior or exterior to a learner, the vicinity of goal and self-regulation gives learners a head start over by enabling them to balance their efforts for moving forward to secure much-desired outcome.

The scores on help-seeking measures were correlated with goal orientation and writing self-regulation. Goal orientation is mediated and invigorated as the result of assuming responsibility which effectuates engagement. Previous studies have also established the linkage between achievement goals and help-seeking (Ryan et al., 1998; Huang, 2011). This suggests that Iranian EFL learners go to any length irrespective of other learners' judgments about their requests for help as to achieve their writing performance objectives, reminding the notion made by Wolters (2010) that help-seeking stands as indicator of the social dimension of self-regulated learning in which learners reach out teachers, parents, peers, and others for the management of their learning. Traditional classroom research shows that high-achieving students with well-developed self-regulatory beliefs and behaviors incline to make use of their teachers and peers as social supports (Zimmerman & Martinez-Pons, 1986).

It is considerable to note that the non-significant correlation between achievement on task and self-regulation in writing is consistent with the findings of the study done by Farsani et al., (2014), likewise. The potential reason for this fact is that self-regulation is more the act of self-adjustment than that of real-task execution. The fact that writing self-regulation was not related to task achievement is because self-regulation is more about the problems in event of writing to be solved than task properties. Self-regulation is more subsidiary to act of writing that helps learners keep the task and self-processes on the right track.

The findings are also incongruent with the study carried out by Kitsantas et al., (2009), evidencing goal orientation not as a significant predictor of students' outcome measures across different subject areas. Noteworthy, the results showed that achievement on task is negatively related to cognition, meta-cognition, motivation, and behavior. The non-existence of a significant link between achievement on task and writing self-regulation could explain the fact that taking a self-regulation inventory is a retrospective act of finding and facing the reasons for failure or triumph in its own place as to prospectively open a venue to success. This is an act of aftermath, surfacing experiences of breakthrough or blockage from which one learns to control future attempts as situations arise, not that of in-context performance in which situated task-specific attempts could be palpable.

This study provides insight in supporting a model of self-regulation in writing in which goal orientations substantially play a significant part. The results corroborate efficiency findings of the factor analytic study by Kaplan et al., (2009) as for structural properties and satisfactory internal consistency and reliability of the integral elements of the inventory for engagement purposes. It is also an extension to the studies that analyze the factors which are key and quintessential to the delicate and detailed process of self-regulation. It endorses that the extension of self-regulated model of learning, developed by (Zimmerman & Kitsantas, 2002), to domain-specific task of writing applied in EFL context is theoretically valid.

The findings give insight into the generalizability of writing self-regulation patterns across cross-cultural contexts. The present study puts premium on the significance of examining self-regulation in writing in EFL settings. The validation of the writing self-regulation model provides a theoretically sound and methodologically valid and reliable measurement for examining writing self-regulation practice of EFL learners. Using this validated scale paves the way for EFL learners to tailor to their needs for transformation as to secure their efficient stays at the learning programs, improve their engagements, and realize their outcome-related socio-educational objectives.

## 6. Conclusion

This study aimed to examine the construct validity and factor structure of writing self-regulation of a sample of Iranian EFL state-run university learners in a domain-specific academic writing context. The findings confirmed that goal, efficacy, and writing self-regulation subcomponents load on a single factor, indicating that they all measure the same construct. The confirmatory factor analysis confirms that the inventory is accredited as a promising tool to gauge writing self-regulation. Although writing self-regulation can be taken as a mechanism to uplift performance, Iranian EFL learners' task achievement was not significantly related to their goal orientations and writing self-regulation. Thus, it can be inferred and concluded that understanding from the purpose and function of self-regulation in relation to achievement on task needs to be renovated.

The results provided valid empirical evidence for the importance of goal orientations in the structure of self-regulation in the writing of Iranian EFL learners. Based on the findings of the study, it can be concluded that self-regulation and goal orientation work hand in hand to direct Iranian EFL learners' writing performances in the classroom. Iranian EFL learners purposefully endorse goals and employ cognitive, meta-cognitive, motivational, and behavioral strategies for self-regulation of their writing tasks. Furthermore, the findings imply that entering EFL learners in goal-conditioned setting gives them a sense of direction in their performance, which ipso facto enhances their spirits to apply self-regulation of writing.

In line with the social cognitive theory, the findings proved that goal-conditioned context is germane to self-regulation in writing of EFL learners. It can be implicated that the goal orientations operate in tandem with the motivational, behavioral, cognitive, and meta-cognitive dimensions in the system of self-regulation in writing. The educational implication construed from the outcomes of this study is that teachers can design goal-based learning contexts which bond a close-knit spot of communion with self-regulation of writing. The findings prompt educators and practitioners to reflect more upon the consequence of goal-conditioned structures for writing.

All told, the validation of the self-regulation in writing instrument in this study can render multiple practical applications to Iranian educational EFL programs. This study paves the way for drawing a picture of self-regulated learning for which a single criterion is not necessarily to be met (Zimmerman & Martinez-Pons, 1988). It permits students to better identify and understand the specific nature of key elements in their patterns of writing performances and weight those for their academic and professional benefits. Educators and curriculum developers can also exploit this measure to gauge motifs of goal orientations to deeper discern learners' motivations and accordingly dispense exigent prescriptions that help them become more efficacious.

The findings also provide empirical implicational evidence on the integrity of the six-factor structure of the writing self-regulation model extrapolated to the EFL educational setting in Iran. These results provided strong support for the reliability and validity of the self-regulation in writing inventory, making it convenient for researchers and teachers to use this invaluable research tool to assess writing self-regulation patterns in university-level English classes in Iranian EFL context. This study corroborates the utility of this inventory for university learners to review their strategically goal-oriented, motivated, behavioral, meta-cognitive, and cognitive self-regulatory behaviors in writing environments.

In a nutshell, the results can be adduced to explain the role of goal orientations in guiding writing self-regulation in EFL contexts. Thus, for elevating learners' self-regulation in writing, it is advised educators to help learners establish goal. There it feels a need for introducing writing self-regulation inventory to learners to acquaint them more with the practice of self-survey of their writing performances as to help them turn into independent learners. The inventory validated in the present study can be exploited by Iranian EFL learners for systematic self-appraisal of their writing progress as to pinpoint the areas of strength and weakness for bestowing on improvement in writing. As witnessed, the findings provided cultural and contextual support for consistency and applicability of a seven-factor model of writing self-regulation, namely, goal, efficacy, meta-cognitive, cognitive, motivational, and behavioral variables in domain-specific context of Iranian EFL writing. The findings, also, proving that the inventory is invariant in the EFL context, provide justifiable evidence for future studies.

## 7. References

- Al-Harthy, I. S., Was, C. A., & Isaacson, R. M. (2010). Goals, efficacy and metacognitive self-regulation a path analysis. *International Journal of Education*, 2(1), 1. <https://doi.org/10.5296/ije.v2i1.357>

- Allal, L. (2018). The co-regulation of writing activities in the classroom. *Journal of Writing Research, 10*(1), 25-60. <https://doi.org/10.17239/jowr-2018.10.01.02>
- Artino, A. R., & Stephens, J. M. (2007). Motivation and self-regulation in online courses: A comparative analysis of undergraduate and graduate students. *Paper presented at the annual meeting of the Association for Educational Communications and Technology, Anaheim, CA. Journal of Computer Assisted Learning, 2008* - Wiley Online Library.
- Bailey, A. (2016). *The Effect of Extended Instruction on Passive Voice, Reduced Relative Clauses, and Modal Would in the Academic Writing of Advanced English Language Learners*. Dissertation and Theses. Paper 3203. DOI: 10.15760/etd.3194
- Bernacki, M. L., Byrnes, J. P., & Cromley, J. G. (2012). The effects of achievement goals and self-regulated learning behaviors on reading comprehension in technology-enhanced learning environments. *Contemporary Educational Psychology, 37*(2), 148-161. <https://doi.org/10.1016/j.cedpsych.2011.12.001>
- Boekaerts, M., Maes, S., & Karoly, P. (2005). Self-regulation across domains of applied psychology: Is there an emerging consensus? *Applied Psychology, 54*(2), 149-154. <https://doi.org/10.1111/j.1464-0597.2005.00201.x>
- Bouffard-Bouchard, T., Parent, S., & Larivee, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students. *International Journal of Behavioral Development, 14*, 153-164. <https://doi.org/10.1177/016502549101400203>
- Canfield, C. & Zastavker, Y. V. (2010). *Achievement Goal Theory: A Framework for Implementing Group Work and Open-Ended Problem Solving*. 40th ASEE/IEEE Frontiers in Education Conference, Washington, DC. <https://doi.org/10.1109/FIE.2010.5673359>
- Cheung, E. M. T. (2008). *How goal orientations and learning environments are related to beliefs in effort-ability relationship*. (Master in Educational Psychology thesis), the University of Hong Kong. [http://dx.doi.org/10.5353%2Fth\\_b4171624](http://dx.doi.org/10.5353%2Fth_b4171624)
- Dinc, Z. F. (2010). Relationship between achievement goal orientation and physical self-perception among students attending physical education teaching. *World Applied Sciences Journal, 11* (6), 662-668.
- Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273. <https://doi.org/10.1037/0033-295X.95.2.256>
- El-Henawy, W. M., Dadour, E.-S. M., Salem, M. M., & El-Bassuony, J. M. (2012). The effectiveness of using self-regulation strategies on developing argumentative writing of EFL prospective teachers. *Journal of the Egyptian Association for Reading and Knowledge, 27*(1), 1-28.
- Elliot, A. J. (1999). Approach and avoidance motivation and achievement goals. *Educational Psychologist, 34*, 149-169. [https://doi.org/10.1207/s15326985ep3403\\_3](https://doi.org/10.1207/s15326985ep3403_3)
- Field, A. (2013). *Discovering statistics using SPSS* (4th ed.). London: SAGE Publications.
- Farsani, M. A., Beikmohammadi, M., & Mohebbi, A. (2014). Self-regulated learning, goal-oriented learning, and academic writing performance of undergraduate Iranian EFL learners. *The Electronic Journal for English as a Second Language, 18* (2), 1-19.
- Flower, L. S., & Hayes, J. R. (1981). *A Cognitive Process Theory of Writing*. *College Composition and Communication, 32*, 365-387. <http://dx.doi.org/10.2307/356600>
- Graham, S., & Harris, K. R. (2000). The role of self-regulation and transcription skills in writing and writing development. *Educational Psychologist, 35*, 3-12. [https://doi.org/10.1207/S15326985EP3501\\_2](https://doi.org/10.1207/S15326985EP3501_2)
- Harris, K., Graham, S., Mason, L., & Friedlander, B. (2008). *Powerful writing strategies for all students*. Baltimore: Paul H. Brookes Publishing Co., Inc

- Harrison, S., & Prain, V. (2009). Self-regulated learning in junior secondary English. *Issues in Educational Research*, 19(3), 2009
- Hayes, J. R. (2012). Modeling and remodeling writing. *Written Communication*, 29, 369–388. <https://doi.org/10.1177/0741088312451260>
- Hidi, S., & Boscolo, P. (2006). Motivation and Writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of Writing Research* (pp. 144-157). New York: The Guilford Press.
- Huang, C. (2011). Achievement goals and achievement emotions: A meta-analysis. *Educational Psychology Review*, 23, 359-388. <https://doi.org/10.1007/s10648-011-9155-x>
- Kaiser, H.F. (1974). An index of factorial simplicity. *Psychometrika*, 39(1), 31-36. <https://doi.org/10.1007/BF02291575>
- Kanlapan, T. C. E., & Velasco, J. C. (2009). Constructing a self-regulation scale contextualized in writing. *TESOL Journal*, 1, 79-94.
- Kaplan, A., Lichtinger, E., & Gorodetsky, M. (2009). Achievement goal orientations and self-regulation in writing: An integrative perspective. *Journal of Educational Psychology*, 101 (1), 51-69. <https://doi.org/10.1037/a0013200>
- Kitsantas, A., Steen, S., & Huie, F. (2009). The role of self-regulated strategies and goal orientation in predicting achievement of elementary school children. *International Electronic Journal of Elementary Education* Vol. 2, Issue 1, October, 2009.
- Kozlowski, S. W. J., & Bell, B. S. (2006) Disentangling achievement orientation and goal setting: Effects on self-regulatory processes. *Journal of Applied Psychology*, 91, 900- 916. <https://doi.org/10.1037/0021-9010.91.4.900>
- Lichtinger, E., Kaplan, A., & Gorodetsky, M. (2006). *Exploring self-regulation in writing among junior high-school students: A simulated recalls study*. Paper presented at the 10th International Conference of the EARLI Special Interest Group on Writing, Antwerp, Belgium.
- Linnenbrink, E. A. (2005). The dilemma of performance-approach goals: the use of multiple goal contexts to promote students' motivation and learning. *Journal of Educational Psychology*, 97(2), 197–213. <https://doi.org/10.1037/0022-0663.97.2.197>
- MacArthur, C. A., Philippakos Z. A., Ianetta, M. (2015). Self-regulated strategy instruction in college developmental writing. *J. Educ. Psychol.* 107 855–867. <https://doi.org/10.1037/edu0000011>
- Magno, C. (2008). Reading strategy, amount of writing, metacognition, metamemory, and apprehension as predictors of English written proficiency. *Asian EFL Journal*, 29, 15–48.
- Malpass, J. R., O'Neil, H. F., & Hocevar, D. (1999). Self-regulation, goal orientation, self-efficacy, worry, and high-stakes math achievement for mathematically gifted high school students", *Roepers Review*, 21(4), 281-297. <https://doi.org/10.1080/02783199909553976>
- Meece, J. L., & Miller, S. D. (1999). Changes in elementary school children's achievement goals for reading and writing: Results of a longitudinal and an intervention in study. *Scientific studies of reading*, 3, 207-229. [https://doi.org/10.1207/s1532799xssr0303\\_2](https://doi.org/10.1207/s1532799xssr0303_2)
- Midgley, C., Maehr, M. L., Hruda, L. Z., Anderman, E., Anderman, L., Freeman, K. E., et al. (2000). *Patterns of Adaptive Learning Survey(PALS)*. Ann Arbor: University of Michigan.
- Nicholls, J. G. (1989). *The competitive ethos and democratic education*. Cambridge, MA: Harvard University Press.
- Pajares, F., Britner, S. L., & Valiante, G. (2000). Writing and science achievement goals of middle school students. *Contemporary Educational Psychology*, 25, 4067422. <https://doi.org/10.1006/ceps.1999.1027>




- Pajares, F., & Cheong, Y. F. (2003). Achievement goal orientations in writing: A developmental perspective. *International Journal of Educational Research*, 39, 437–455. <https://doi.org/10.1016/j.ijer.2004.06.008>
- Pintrich, P. R., & Garcia, T. (1991). Student goal orientation and self-regulation in the college classroom. In M. L. Maehr and P. R. Pintrich (Eds.), *Advances in Motivation and Achievement* (pp. 371-402). Greenwich, CT: JAI Press
- Pintrich, P. R. & Smith, D. A. F., Garcia, T. & McKeachie, W. J. (1991). *A Manual for the Use of the Motivated Strategies for Learning Questionnaire (MSLQ)*. Ann Arbor, MI: NCRIPAL, The University of Michigan, 1991: 3.
- Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and Psychological Measurement*, 53, 801–813. <https://doi.org/10.1177/001316449305300302>
- Pratontep, C., & Chinwonno, A. (2008). Self-regulated learning by Thai university students in an EFL extensive reading program. *MANUSYA: Journal of Humanities*, 11(2), 104-124. <https://doi.org/10.1163/26659077-01102006>
- Prior, P. (2006). A sociocultural theory of writing. In C. A. MacArthur, S. Graham, & J. Fitzgerald (Eds.), *Handbook of writing research* (pp. 54-66). New York: Guilford Press.
- Ruan, Z. (2005). A metacognitive perspective on the growth of self-regulated EFL student writers. *Reading Working Papers in Linguistics*, 8, pp.175-202.
- Ryan, A. M., Gheen, M. H., & Midgley, C. (1998). Why do some students avoid asking for help? An examination of the interplay among students' academic efficacy, teachers' social-emotional role, and the classroom goal structure. *Journal of Educational Psychology*, 90(3), 528-535. <https://doi.org/10.1037/0022-0663.90.3.528>
- Sala-Bubare, A & Castello, M. (2018). Writing regulation processes in higher education: A review of two decades of empirical research. *Reading and Writing*, 31(4), 757-777. <https://doi.org/10.1007/s11145-017-9808-3>
- Schunk, D. H. & Zimmerman, B. J. (2007). Influencing Children's Self-Efficacy and Self-Regulation of Reading and Writing through Modeling. *Reading and Writing Quarterly*, 23, 7-25. <https://doi.org/10.1080/10573560600837578>
- Shaw, S., & Falvey, P. (2008). *The IELTS writing assessment revision project: Towards a revised rating scale*. Cambridge, United Kingdom: University of Cambridge ESOL Examinations.
- Syafitri, Kurnia. (2019). *The Correlation between Students' Self Regulation and Their Writing Skill of The Eleventh Grade at State Senior High School 1 Tambusai Utara*. Pekanbaru: Universitas Sultan Syarif Kasim Riau.
- Tavşancıl, E. (2014). *(Attitude Measurement and Data Analysis via SPSS) Tutumların Ölçülmesi ve SPSS İle Veri Analizi*. Nobel Yayıncılık. Ankara. 48.
- Teng, L. S., & Zhang, L. J. (2016). Fostering strategic learning: The development and validation of the writing strategies for motivational regulation questionnaire (WSMRQ). *Asia-Pacific Education Researcher*, 25, 123–134. <https://doi.org/10.1007/s40299-015-0243-4>
- Tseng, W. T., Dörnyei, Z., & Schmitt, N. (2006). A new approach to assessing strategic learning: the case of self-regulation in vocabulary acquisition. *Applied Linguistics*, 78-102. <https://doi.org/10.1093/applin/ami046>
- University of Cambridge: *ESOL Examinations* (2004). *Quick Placement Test* (Version 1). Oxford University Press.
- Uysal, H. H. (2009). A critical review of the IELTS writing test. *ELT Journal*, 64(3), 314–320. <https://doi.org/10.1093/elt/ccp026>
- Wolters, C. (2010). *Self-regulated learning and the 21<sup>st</sup> Century Competencies*. Retrieved July 7, 2011, from <http://www.hewlett.org/pdf>

- Winne, P. H., Jamieson-Noel, D. L., & Muis, K. (2002). Methodological issues and advances in researching tactics, strategies, and self-regulated learning. In Pintrich, P.R. & Maehr, M. L. (Eds.), *Advances in motivation and achievement: New directions in measures and methods* (Vol. 12, pp. 121–155). Greenwich, CT: JAI Press.
- Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American Educational Research Journal*, 31(4), 845-862. <https://doi.org/10.3102/00028312031004845>
- Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology*, 94, 660-668. <https://doi.org/10.1037/0022-0663.94.4.660>
- Zimmerman, B.J., & Kitsantas, A. (2007). Reliability and validity of self-efficacy for Learning Form (SELF) scores of college students. *Journal of Psychology*, 215, 157-163. <https://doi.org/10.1027/0044-3409.215.3.157>
- Zimmerman, B. J., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal*, 23: 614-28. <https://doi.org/10.3102/00028312023004614>
- Zimmerman, B., & Martinez-Pons, M. (1988). Construct validation of a strategy model of student self-regulated learning. *Journal of Educational Psychology*, 80, 284-290. <https://doi.org/10.1037/0022-0663.80.3.284>
- Zimmerman, B. J., & Schunk, D. H. (Eds.). (1989). *Self-regulated learning and academic achievement: Theory, research, and practice*. New York: Springer-Verlag. [https://doi.org/10.1007/978-1-4612-3618-4\\_1](https://doi.org/10.1007/978-1-4612-3618-4_1)
- Zumbrunn, S. K., Marrs, S. A., & Mewborn, C. (2016). Toward a better understanding of student perceptions of written feedback: a mixed methods study. *Reading and Writing*, 29(2), 349–370. <https://doi.org/10.1007/s11145-015-9599-3>



<https://doi.org/10.22077/ali.2024.7176.1023>

# A Comparative Analysis of the Impact of Two Pedagogical Approaches on the Production of Idiomatic Expressions among EFL Learners

Mehdi Mehranirad<sup>1</sup>   
 Reyhaneh Garazhian<sup>2</sup>   
 Reza Zabihi<sup>3\*</sup> 

<sup>1</sup>Associate Professor in Applied Linguistics, Department of English, Faculty of Letters and Humanities, University of Neyshabur, Neyshabur, Iran

<sup>2</sup>M.A. in English Language Teaching, Department of English, Faculty of Letters and Humanities, University of Neyshabur, Neyshabur, Iran

<sup>3\*</sup>Associate Professor in Applied Linguistics, Department of English, Faculty of Letters and Humanities, University of Neyshabur, Neyshabur, Iran

## ABSTRACT

The ongoing debate surrounding the effectiveness of instructional models, particularly within the context of English as a Foreign Language (EFL) education, underscores the need for empirical investigation. This study sought to address this discourse by examining the differential impact of two pedagogical approaches on the production of idiomatic expressions. Employing a quasi-experimental design, the research compared the performance outcomes of two groups of learners, one exposed to deductive instruction and the other to inductive instruction. The participants engaged in a comprehensive exploration of 64 English idiomatic expressions over the course of the study. Data collection included a pre-test, eight immediate post-tests, a delayed post-test, and an open-ended questionnaire. The 20-session experiment, integrated seamlessly into the regular classroom instruction, involved the alternation of deductive and inductive approaches in teaching eight idiomatic expressions per session. Findings revealed significant improvements in both groups' scores, yet a noteworthy distinction emerged in favor of the inductive instruction. This outcome implies that an inductive pedagogical approach may exert a significant influence on the production of idiomatic expressions among EFL learners. The study contributes insights into the ongoing pedagogical discourse and underscores the potential benefits of adopting inductive strategies in the teaching of idiomatic expressions.

**KEYWORDS:** Deductive teaching; Idiomatic expressions; Inductive teaching; Memory retention

## ARTICLE HISTORY

Received: 15 April 2023  
 Revised: 06 June 2023  
 Accepted: 21 August 2023  
 Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [mehrani@neyshabur.ac.ir](mailto:mehrani@neyshabur.ac.ir)

## 1. Introduction

The acquisition of idiomatic expressions presents a distinctive challenge for second language (L2) learners due to their intricate nature. Despite their significance for achieving native-like fluency (Wray, 2000), idiomatic expressions often find themselves relegated in English course books, with language teachers often opting to overlook their instruction. McLaughlin et al, (1983)



contends that the use of idiomatic expressions is a distinguishing feature between native and non-native speakers, highlighting the inherent complexity faced by the latter. Consequently, while the acquisition of idiomatic expressions poses a formidable challenge, it is imperative for course book designers to allocate due attention to their inclusion into syllabi, and educators to employ the most effective teaching methodologies.

Presently, consensus exists regarding the importance of teaching idiomatic expressions to L2 learners. However, when it comes to the discussion of the most effective approach for teaching idiomatic expressions, educators' opinions diverge. In fact, this divergence of perspectives often manifests as a debate between deductive and inductive teaching methods. Deductive teaching involves imparting learners with explicit rules and subsequently applying those rules to meaningful examples (Thornbury, 1999). Conversely, inductive teaching methods provide learners with examples, prompting them to analyze and infer the underlying rules independently (Ellis, 2010). Despite numerous studies exploring the efficacy of deductive and inductive methods for teaching grammar, there seems to be a lack of conclusive consensus among researchers, particularly on the impact of these methods on the production of idiomatic expressions.

To address this gap, the current study was set to investigate the relative influence of deductive and inductive teaching methods on the production of English idiomatic expressions among Iranian EFL learners. While the relative advantages and disadvantages of these methods have been extensively explored in English-dominant societies, the generalizability of their findings to regions where English is a foreign language remains uncertain. The lack of consensus is evident in studies favoring either inductive (e.g. Abdolmanafi Rokni, 2009) or deductive approaches (Berendse, 2012; Erlam, 2003; Nazari, 2012; Shih, 2008). Additionally, there is a scarcity of research comparing these methods in the specific context of ESL/EFL classes. Therefore, this study aims to scrutinize deductive and inductive teaching methods in the Iranian EFL context, assessing their impact on English idiomatic expression production, learner attitudes towards each method, and the challenges faced by learners. On this basis, this study addresses the following research questions.

1. Do different teaching methods, specifically deductive versus inductive approaches, yield variations in the production of taught L2 idiomatic expressions?
2. What are the students' attitudes toward the effectiveness of each instructional method?
3. What challenges do Iranian EFL students encounter in the acquisition of idiomatic expressions?

Aligned with Doughty and Williams (1998), who advocate for the efficacy of the inductive approach in language teaching, and also based on Mohamadi Asl's (2013) findings emphasizing the effectiveness of teaching idiomatic expressions in extended contexts, we hypothesized that students instructed inductively would demonstrate greater accuracy in the production of English idiomatic expressions.

## 2. Literature Review

Idiomatic expressions, characterized by their intricate syntax and meanings modifying literal interpretations, present a formidable challenge for second language (L2) learners. Despite the universal acknowledgment of their significance in achieving native-like fluency, there is no consensus over a universally accepted definition for idiomatic expressions. Nattinger and De Carrico's (1992) description of idiomatic expressions as "complex bits of frozen syntax" reflects the intricacies faced by learners in comprehending these linguistic constructs. Existing research on L2 idiomatic expression learning reveals a diverse range of findings (Irujo, 1986; Cooper, 1999; Alhaysony, 2017).

Early investigations by Irujo (1986) delved into the organizational aspects of L2 idiomatic expressions. His study explored the influence of the first language on advanced learners' comprehension and production of English idiomatic expressions. Notably, the study categorized idiomatic expressions based on their similarity to learners' native language (Spanish). Findings highlighted the ease of understanding and production for idiomatic expressions identical to those in the native language, suggesting a nuanced relationship between L1 and L2 idiomatic expression acquisition.

Other studies have revealed that the age of learners can be a pivotal factor influencing their acquisition of idiomatic expressions. In fact, research indicates that individuals under the age of nine tend to interpret idiomatic expressions literally (Cacciari & Levarato, 1989). This phenomenon, supported by several studies (e.g. Cacciari & Levarato, 1989; Douglas & Peel, 1979), is supported by Piaget's theory of cognitive development, attributing the literal interpretation tendency to the limited abstract thinking abilities of younger children (Piaget, 1959).

Contextual influence on learning various idiomatic expressions is also well-documented. For instance, Mohamadi Asl (2013) explored the role of context in the learning experiences of Iranian ESL students, revealing that extended contextual exposure significantly improved immediate and delayed post-test scores compared to limited and decontextualized contexts. Consistent with this finding, other studies emphasize the facilitating effect of context on the acquisition of idiomatic expressions across various age groups (Cacciari & Levarato, 1989).

Despite the extensive body of literature various pedagogical instructions, specifically deductive versus inductive approaches, limited research has investigated the relative effectiveness of these teaching methods on learning idiomatic expressions. As a matter of fact, most of the previous studies have exclusively focused on the acquisition syntactic elements. For example, Erlam (2003) examined the impact of deductive and inductive teaching methods on the acquisition of direct object pronouns in French as a second language. The deductive group, receiving explicit rule presentations, exhibited significant improvement compared to the inductive group, emphasizing the potential efficacy of deductive approaches. More recently, Berendse's (2012) study on English past and present tenses contributed to the debate, revealing an initial similarity between the short-term effect of deductive and inductive teaching but demonstrating the long-term superiority of deductive instruction. This discrepancy in findings highlights the complexity of the deductive-inductive dichotomy and its context-dependent outcomes.

Nevertheless, studies by Haight, Herron, and Cole (2007), Abdolmanafi Rokni (2009), and Vogel et al. (2011) challenge the presumed superiority of deductive methods. Haight et al.'s (2007) research on elementary learners of French as a foreign language revealed higher mean scores for inductive instruction across both short-term and long-term tests. Abdolmanafi Rokni (2009) similarly reported the superiority of inductive approaches among Iranian university EFL students. Vogel et al.'s (2011) findings, while emphasizing the short-term efficacy of inductive approaches, underscored inconclusive long-term effects.

Studies such as those conducted by Jean and Simard (2013), Chalipa (2013), and Lee and Lin (2019) present a nuanced perspective, revealing the similarity of deductive and inductive methods in terms of learning outcomes. More specifically, Jean and Simard's (2013) empirical study revealed comparable learning outcomes between deductive and inductive approaches, despite student preferences for deductive instruction. In the Iranian context, Chalipa (2013) found no significant difference in the influence of inductive and deductive grammar teaching on Iranian university students, challenging some of the previous studies that presumed a greater efficacy for deductive approaches. Similarly, Lee and Lin's (2019) investigation into the acquisition and retention of vocabulary through inductive and deductive teaching methods found no significant differences between groups, challenging the prevailing assumptions about the efficacy of one approach over the other.

Alhaysony's (2017) observation that language teachers traditionally prioritize grammar over idiomatic expressions due to the perceived difficulty in teaching them highlights the need for dedicated attention to idiomatic expression instruction. Pimenova's (2012) identification of five main challenges in learning idiomatic expressions, including unknown vocabulary, cultural differences, and lack of contextualization, further emphasizes the multifaceted nature of idiomatic expression acquisition.

In conclusion, the literature reflects numerous studies investigating the role of various factors related to the acquisition of idiomatic expressions. There are also a considerable number of empirical studies, comparing the effects of deductive and inductive instruction on the development of language learners' grammatical competence. However, there is a scarcity of research comparing the actual impact of these methods for learning idiomatic expressions in the specific context of ESL/EFL classes. The varied and inconclusive findings across studies underscore the need for further exploration, particularly in the Iranian EFL context, to develop more effective guidelines for teaching idiomatic expressions. Therefore, the present study aims to bridge this knowledge gap by examining the impact of deductive and inductive teaching methods on the production of English idiomatic expressions among Iranian EFL learners, considering learner attitudes and challenges encountered in the process.

### 3. Methodology

#### 3.1. Participants

The study included a total of 32 participants; however, due to attrition, only 26 participants completed the study. The participants consisted of 16 males and 10 females, aged 13 to 17 ( $M = 15.3$ ,  $SD = 0.95$ ). These Iranian non-native English learners had a maximum of one year of experience in learning English as a foreign language. They were monolingual native speakers of Persian and were studying English in a private language institute in Neyshabur during the spring semester of 2021.

Participants were enrolled in four intact classes, and at the beginning of the course, they were informed about additional activities focusing on idiomatic expressions. Participation in these activities was voluntary, with no mandatory requirement, and no additional marks were assigned for participation. Homogeneity in language skills was established through the Face-to-face Cambridge University placement test, which categorized all participants as elementary learners. Two classes were randomly assigned to receive deductive instruction, while the other two received inductive instruction. Both groups received instruction twice a week for an hour and a half over a 10-week period.

- Group A: 12 students (7 male, 5 female), aged 13 to 17 ( $M = 15.08$ ,  $SD = 1.08$ ), received deductive teaching method for English idiomatic expressions.

- Group B: 14 students (9 male, 5 female), aged 13 to 17 ( $M = 15.14$ ,  $SD = 1.02$ ), received inductive teaching method for English idiomatic expressions.

### 3.2. Instruments

The study utilized various instruments, including a placement test, a pre-test, post-tests, and a questionnaire, all specifically designed for this research. Initially, Cambridge University placement test was employed to assess the homogeneity of learners. This test included questions resembling natural conversation interactions, aiding in placing students in appropriate classes based on their language proficiency.

To measure the students' baseline knowledge of idiomatic expressions, a pre-test consisting of 24 completion questions was administered. The questions presented various scenarios requiring appropriate idiomatic expressions. The pre-test results indicated a lack of familiarity with idiomatic expressions among students.

Eight immediate post-tests, administered at the end of each session, comprised eight completion questions assessing learners' understanding of idiomatic expressions. Additionally, a delayed post-test, consisting of 24 completion questions, was administered at the end of the course.

Finally, an open-ended questionnaire specifically designed for the purpose of this study was administered during the final session of the course to evaluate students' attitudes towards deductive and inductive teaching methods for idiomatic expressions and identify the difficulties they encountered. In an attempt to eliminate potential language barriers, the questionnaire was designed in Persian. The data collected through the questionnaire were thematically analyzed in order to answer the second and the third research questions.

### 3.3. Procedure

Participants were divided into two groups based on the treatment received – deductive or inductive teaching methods. At the outset of the course, the pre-test was used to assess participants' idiomatic expression knowledge. The deductive group (12 students) received explanations in class with opportunities for questions and direct feedback on errors. Activities included completing handouts and sentence writing for explanation. For instance, the idiomatic expression "As simple as ABC" was introduced with meaning and examples, followed by substitution exercises. In the inductive group (14 students), sentences with idiomatic expressions were presented, and students worked in groups to deduce meanings, with the correct meanings later presented by the professor. Both groups received 30 hours of instruction (3 hours weekly) over 10 weeks. The post-tests were then used, both at the end of each session and at the end of the course to assess students' improvement.

## 4. Results and discussion

To assess the baseline knowledge of the participants' idiomatic expressions, an independent samples t-test was conducted, comparing the pre-test scores of both the deductive and inductive groups. The results indicated no statistically significant difference in the knowledge of target idiomatic expressions between the deductive and inductive groups before the treatment  $t(24) = 2.579$ ,  $p = .121$ .

Having ensured that students in both groups were not different in terms of the knowledge of idiomatic expressions prior to the instruction, we then conducted a series of comparisons across eight immediate post-tests administered at the end of each treatment session. The tests, each consisting of eight items, were designed to maintain consistent difficulty levels using the same rubrics and question stems. Conducting independent samples t-tests, the results, summarized in Table 1, largely indicated no statistically significant difference between the deductive and inductive groups in terms of immediate retention of idiomatic expressions.

**Table 1.** Results of the scores of the immediate post-tests in the deductive and inductive groups

Immediate Post tests	Deductive group		Inductive group		<i>p</i>
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	
1 <sup>st</sup> test	5.91	1.31	7.00	1.24	0.04
2 <sup>nd</sup> test	6.66	1.15	7.14	1.29	0.33
3 <sup>rd</sup> test	6.58	1.67	6.71	1.72	0.84
4 <sup>th</sup> test	6.00	1.85	6.50	2.21	0.54
5 <sup>th</sup> test	5.66	1.72	6.21	1.57	0.40

6 <sup>th</sup> test	5.33	1.55	6.35	1.64	0.11
7 <sup>th</sup> test	6.25	1.65	5.85	2.03	0.59
8 <sup>th</sup> test	5.58	1.78	6.35	1.15	0.19

Subsequent analyses included the comparison of the initial and final performances of the students in each group. As presented in Table 2, a paired-samples t-test was employed to examine the impact of the inductive teaching method on delayed idiomatic expressions test scores. The results reveal a statistically significant difference  $t(13) = -6.98, p = 0.00$  in scores between the pre-test ( $M = 6.14, SD = 2.34$ ) and the post-test ( $M = 12.64, SD = 3.62$ ) in the inductive group, indicating a significant increase in post-treatment scores.

To assess the effects of the deductive teaching method on delayed idiomatic expressions test scores, a paired-samples t-test was similarly conducted. As illustrated in Table 2, statistical analyses revealed a significant difference,  $t(11) = -2.70, p = 0.02$  between the pre-test scores ( $M = 5.33, SD = 1.30$ ) and post-test scores ( $M = 8.75, SD = 4.20$ ) in the deductive group.

**Table 2.** The descriptive statistics of pre-test and delayed post-test for the both groups

Tests/groups		N	Mean	SD	Min	Max
Pre-test	Deductive	12	5.33	1.30	3	8
	Inductive	14	6.14	2.34	2	11
Post-test	Deductive	12	8.75	4.20	5	17
	Inductive	14	12.64	3.62	3	19

Finally, to examine the disparity in delayed post-test scores between the deductive and inductive groups, an independent samples t-test was executed. A statistically significant difference emerged in favor of the inductive group  $t(24) = -2.53, p = 0.018$ , indicating that the effect of inductive instruction was statistically more significant than the effect of deductive instruction.

The second research question delved into students’ preferences regarding teaching methods, elicited through a questionnaire. Participants predominantly favored the inductive teaching approach. Although some acknowledged its inherent difficulty, they underscored its efficacy in facilitating better retention of idiomatic expressions. An illustrative response from an inductive course participant reveals this sentiment.

At first, it was really difficult to guess the meaning [of an idiomatic expression]. But gradually it got easier. It took a long time and maybe we were confused at times, sometimes I thought that I would not remember a single word, but when I saw the questions, I remembered all our discussions.

In justifying their choice, students recurrently highlighted the effectiveness of guessing activities. The perceived advantage lay in leveraging collective insights during discussions with their peers to come up with the meaning of the expressions, aiding better recall during assessments. This resonates with findings by Koschmann et al. (1996), emphasizing the cognitive benefits of discovery learning and collaborating with peers in fostering deep reflections and the incorporation of new ideas into existing knowledge.

Moreover, students expressed a reliance on contextual cues for memorizing idiomatic expressions, reinforcing the notion that a rich context enhances learning. While students might resort to their native language when context is absent, the preference is inclined towards utilizing a rich context for learning L2 idiomatic expressions.

In synthesis, the empirical evidence highlights the fact that understanding the meaning of single-word items does not necessarily translate to automatic comprehension of multi-word items (Cacciari and Levarato, 1989; Mohamadi Asl, 2013; Pimenova, 2012). Consequently, students posit that the inductive teaching method, fostering collaborative discussions within a rich contextual framework, renders the production of idiomatic expressions more accessible and accurate.

The third research question in this study aimed to discern the challenges learners face while acquiring idiomatic expressions. As explained earlier, the researchers used a survey questionnaire to elicit data pertinent to this research question. The survey, already used in some previous studies (Alhaysony, 2017; Cooper, 1999; Irujo, 1986; Lontas, 1999; Pimenova, 2012; Saleh & Zakaria, 2013), comprised seven items. Students were asked to select the items that posed difficulties in their acquisition of idiomatic expressions.

Table 3 encapsulates the outcomes, revealing that a substantial number of learners experienced various problems with comprehending idiomatic expressions. The consensus among students was that combining vocabulary meanings within an

idiomatic expression did not necessarily unveil its intended meaning; at times, it even ‘mised’ learners from the actual connotation. A student’s testimonial illustrates this predicament.

Sometimes the vocabulary in the idiomatic expression causes some problems. For example, there was an idiomatic expression “A pain in the neck”, and we had a unit about health, and we learned about some health problems such as a stomachache. The teacher told us that “a stomachache is when you have a pain in your stomach”. Then when I saw this idiomatic expression, I thought that this was related to some health problem which I do not know the exact word yet.

**Table 3.** Difficulties learners face during learning idiomatic expressions

Challenges experienced by learners	Frequency
The difference between the figurative and the literal meanings of idiomatic expressions.	9
Lack of enough contextual clues.	7
Lack of equivalence in Persian.	3
The interference from the first language.	2
Lack of experience and exposure with idiomatic expressions.	2
Lack of cultural background.	2
They often contain unfamiliar words.	1

Additionally, the findings underscored the complexity associated with idioms, particularly due to the absence of dedicated courses focused on idioms. Participants articulated that idiomatic expressions are often ‘marginalized’ within their English courses. This was particularly explained by one of the students in the following way.

I love idiomatic expressions. When I was in elementary school, I had a teacher who taught us some idiomatic expressions by telling their stories. But from then on, there was not a time when we could talk about them. And now it is really good that in this class we have some time specific to idiomatic expressions.

A further potential difficulty in acquiring idiomatic expressions stems from the challenge of understanding expressions that lack counterparts in the learners’ native language, Persian. Students’ reflections shed light on this point.

Those idiomatic expressions which have analogues in Persian, e.g., ‘Roll up your sleeves’ or ‘A wolf in sheep’s clothing,’ are much easier to understand and even to use in a sentence. But those that we do not have in Persian, e.g., ‘Once in a blue moon,’ although it is not really hard, because we do not have any analogue in Persian, remembering them seems a bit hard.

This observation aligns with Pimenova’s (2012) findings, highlighting that L2 idiomatic expressions lacking counterparts in the mother tongue pose greater difficulty in comprehension and usage.

Beyond the aforementioned challenges, participants’ responses highlighted that idiomatic expressions featuring unfamiliar words and lacking contextual cues presented the most formidable challenges in terms of comprehension and production. Participants emphasized that context helps them get closer to meaning. An illustrative example from a participant emphasized the role of context.

Some idiomatic expressions are so much alike in my eye. For example, ‘Money doesn’t grow on the tree’ and ‘Cost an arm and a leg’ are very similar to each other. And in the test, I could differentiate between these two with the help of context.”

This finding aligns with previous studies demonstrating how a supportive context facilitates the understanding of idiomatic expressions (Mohamadi Asl, 2013; Irujo, 1986; Rohani et al., 2012).

Another reported difficulty seemed to be rooted in the interference caused by the learners’ native language, Persian. Some learners encountered challenges when the idiomatic expression’s meaning seemed unrelated to its constituent vocabulary. For instance:

"The meaning of some idiomatic expression has got nothing to do with its words. When the teacher told us the meaning of ‘it is raining cats and dogs,’ I always forget the meaning and I think it means that there are so many cats and dogs somewhere, because in Persian we have an idiomatic expression which is a little like this ‘از زمین و آسمان می بارد .....’ and it means that there are many objects somewhere. But if I could practice with some examples, then I would be able to remember the situation and would not make mistakes anymore.

This difficulty is consistent with Derakhshan and Karimi’s (2015) findings, which highlighted the interference of the first language with the second language, influenced by factors such as prior knowledge, language proficiency, and structural differences between the two languages.

## 5. Conclusion

This study aimed to investigate the impact of different teaching methods on the production of idiomatic expressions, students' preferences for these methods, and the challenges faced during the process of learning idiomatic expressions. In addressing the first research question, immediate post-tests were employed to evaluate the impact of inductive and deductive teaching methods on the accuracy of written production. The tests, administered after each session, revealed no significant differences between the two groups. This absence of a significant difference is consistent with findings from previous studies (e.g. Allison, 1959; Chalipa, 2013; Dotson, 2010; Jean & Simard, 2013). Chalipa (2013) reported similar results in terms of short-term learning, but in the current study, the inductive method yielded significantly superior results in long-term learning. This difference was particularly more pronounced in the inductive group. In fact, the inductive group displayed a more substantial increase in average scores of the delayed post-test. The inferential statistics underscored the significance of this difference, suggesting a more pronounced benefit from the inductive method in long-term retention.

The second research question explored participants' preferences for different teaching methods. The majority of participants expressed a preference for the inductive approach over the deductive one. This finding contradicts some previous studies (e.g. Dotson, 2010; Jean & Simard, 2013; Vogel et al., 2011). Participants believe that group work fosters collaborative learning, enabling them to assist each other in grasping language-specific nuances and resolving uncertainties in English.

The third research question delved into the challenges faced by learners in understanding and producing idiomatic expressions. The results unveiled several difficulties, including the absence of a cultural background for many idiomatic expressions. Iranian EFL learners encountered challenges in learning L2 idiomatic expressions without L1 analogs. Additionally, the study highlighted that idiomatic expressions are often neglected in the EFL curriculum, necessitating increased attention in course design. The significance of context in learning idioms emerged as a pivotal factor, supporting findings from previous research (Pimenova, 2012; Saleh & Zakaria, 2013). Recognition of the structural intricacies and comprehension of idiomatic expressions demand enhanced focus in EFL courses.

In conclusion, this study offers valuable insights into the nuanced impact of teaching methods on idiom acquisition, learners' preferences, and the inherent difficulties faced. The findings emphasize the need for pedagogical adaptations to address these challenges and underscore the importance of integrating idiomatic expressions into EFL curricula.

## 6. References

- Abdolmanafi Rokni, S. J. (2009). Explicit grammar instruction. *Language in India*, 9, 117-128.
- Alhaysony, M. H. (2017). Strategies and Difficulties of Understanding English Idioms: A Case Study of Saudi University EFL Students. *International Journal of English Linguistics*, 7(3), 70. <https://doi.org/10.5539/IJEL.V7N3P70>
- Allison, L. W. (1959). Inductive vs. deductive method: Comparative effectiveness in teaching English grammar to general groups of high school seniors. (Doctoral dissertation). ProQuest Dissertations and Theses.
- Berendse, E. (2012). A comparison between the effectiveness of inductive and deductive instruction in L2 English classroom in a L1 Dutch environment. (Unpublished master's thesis). Language and Culture, Utrecht University.
- Cacciari, C., & Levorato, M. C. (1989). How children understand idioms in discourse. *Journal of Child Language*, 16(2), 387–405. <https://doi.org/10.1017/S0305000900010473>
- Chalipa, S. (2013). The effect of inductive vs. deductive instructional approach in grammar learning of ESL learners. *Deccan International Journal of Advanced Research Studies*, 1(1), 76-90.
- Cooper, T. C. (1998). Teaching idioms. *Foreign Language Annals*, 31(2), 255–266. <https://doi.org/10.1111/j.1944-9720.1998.tb00572.x>
- Derakhshan, A., & Karimi, E. (2015). The interference of first language and second language acquisition. *Theory and Practice in Language Studies*, 5(10), 2112. <https://doi.org/10.17507/tpls.0510.19>
- Dotson, E. (2010). The effects of deductive and guided inductive approaches on the learning of grammar in an advanced college French course. Dissertation Abstracts International, A: The Humanities and Social Sciences, 4374.
- Doughty, C., & Williams, J. (1998). *Focus on form in classroom second language acquisition*. New York: Cambridge University Press.

- Douglas, J. D., & Peel, B. (1979). The development of metaphor and proverb translation in children grades 1 through 7. *The Journal of Educational Research*, 73(2), 116–119. <https://doi.org/10.1080/00220671.1979.10885218>
- Ellis, R. (2008). Investigating grammatical difficulty in second language learning: Implications for second language acquisition research and language testing. *International Journal of Applied Linguistics*, 18(1), 4–22. <https://doi.org/10.1111/j.1473-4192.2008.00184.x>
- Erlam, R. (2003). The effects of deductive and inductive instruction on the acquisition of direct object pronouns in French as a second language. *The Modern Language Journal*, 87(2), 242–260. <https://doi.org/10.1111/1540-4781.00188>
- Haight, C. E., Herron, C., & Cole, S. P. (2007). The effects of deductive and guided inductive instructional approaches on the learning of grammar in the elementary foreign language college classroom. *Foreign Language Annals*, 40(2), 288–310. <https://doi.org/10.1111/j.1944-9720.2007.tb03202.x>
- Irujo, S. (1986). Don't put your leg in your mouth: Transfer in the acquisition of idioms in a second language. *TESOL Quarterly*, 20(2), 287. <https://doi.org/10.2307/3586545>
- Jean, G., & Simard, D. (2013). Deductive versus inductive grammar instruction: Investigating possible relationships between gains, preferences, and learning styles. *System*, 41(4), 1023–1042. <https://doi.org/10.1016/j.system.2013.10.008>
- Koschmann, T., Kelson, A. C., Feltovich, P. J., & Barrows, H. S. (1996). Computer-supported problem-based learning: A principled approach to the use of computers in collaborative learning. In T. Koschmann (Ed.), *Computer-supported collaborative learning: Theory and practice of an emerging paradigm* (pp. 83–124). Mahwah, New Jersey: Lawrence Erlbaum.
- Lee, P., & Lin, H. (2019). The effect of the inductive and deductive data-driven learning (DDL) on vocabulary acquisition and retention. *System*, 81, 14–25. <https://doi.org/10.1016/j.system.2018.12.011>
- Liontas, J. (2017). Why teach idioms? A challenge to the profession. *Iranian Journal of Language Teaching Research*, 5(3), 5-25.
- Liontas, J. I. (1999). *Developing a pragmatic methodology of idiomaticity: The comprehension and interpretation of SL vivid phrasal idioms during reading* (Doctoral dissertation, The University of Arizona).
- Liontas, J. I. (2003). Killing two birds with one stone: Understanding Spanish VP idioms in and out of context. *Hispania*, 86(2), 289. doi: <https://doi.org/10.2307/20062862>
- McLaughlin, B., White, D., McDevitt, T., & Raskin, R. (1983). Mothers' and fathers' speech to their young children: similar or different? *Journal of Child Language*, 10(1), 245–252. <https://doi.org/10.1017/S0305000900005286>
- Mohamadi Asl, F. (2013). The impact of context on learning idioms in EFL classes. *MEXTESOL Journal*, 37(1), 1-12.
- Nattinger, M. and S. De Carrico. (1992). *Lexical phrases and language teaching*. London. Oxford University Press.
- Nazari, N. (2012). The effect of implicit and explicit grammar instruction on learners' achievements in receptive and productive modes. *Procedia - Social and Behavioral Sciences*, 70, 156-162. <https://doi.org/10.1016/j.sbspro.2013.01.051>
- Piaget, J. (1959). *The language and thought of the child*. Routledge & Paul.
- Pimenova, N. (2012). *Idiom comprehension strategies used by English and Russian language learners in a think-aloud study*. Biblio Bazaar.
- Rohani, G., Ketabi, S., & Tavakoli, M. (2012). The effect of context on the EFL learners' idiom processing strategies. *English Language Teaching*, 5(9).
- Saleh, N., & Zakaria, N. (2013). Investigating the difficulties faced in understanding, and strategies used in processing English idiom by the Libyan students. *International Journal of English Language and Translation Studies*, 1(2), 69-90.

- Shih, J. Y. (2008). Effects of inductive and deductive approaches in grammar instruction for junior high school students of differing English proficiency (Unpublished master's thesis). National Pingtung Institute of Commerce, Taiwan.
- Thornbury, S. (1999). *How to teach grammar*. Pearson Education.
- Vogel, S., Herron, C., Cole, S. P., & York, H. (2011). Effectiveness of a guided inductive versus a deductive approach on the learning of grammar in the intermediate-level college French classroom. *Foreign Language Annals*, 44(2), 353–380. <https://doi.org/10.1111/j.1944-9720.2011.01133.x>
- Wray, A. (2000). Formulaic sequences in second language teaching: Principle and practice. *Applied Linguistics*, 21, 463–489. <https://doi.org/10.1093/applin/21.4.463>





<https://doi.org/10.22077/ali.2023.7239.1026>

# Wh-acquisition: A Cross-linguistic Comparison of Persian and English

Masoomeh Taghizadeh<sup>1</sup>   
Somayeh Baniasad-Azad<sup>2\*</sup>

<sup>1</sup>Ph.D. in Applied Linguistics, Department of English Language & Literature, Yazd University, Yazd, Iran.

<sup>2\*</sup>Assistant Professor in Applied Linguistics, Department of Foreign Languages, Shahid Bahonar University, Kerman, Iran

## ARTICLE HISTORY

Received: 16 March 2023  
Revised: 01 June 2023  
Accepted: 13 August 2023  
Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [s.baniasadazad@uk.ac.ir](mailto:s.baniasadazad@uk.ac.ir)

## ABSTRACT

This research aimed to provide insights into L1 acquisition with respect to the formation of Wh-questions by a Persian and an English child. Similarities and differences concerning the variety of patterns produced (both adult-like and non-adult-like) and the frequency and emergence order of each pattern/ Wh-word were taken into consideration. Based on the findings, it is argued that the Persian grammar licenses a wider range of Wh-questions patterns. Besides, unlike the prevalent in-situ state of the Wh-structure in Persian, the present study indicated it to be mostly Wh-fronted in the data production of the child at the specified developmental stages. Moreover, the English child exhibited more instances of non-adult production which can be justified by the higher complexity of the structure in English, as it demands both dislocation of the Wh-word and inversion. The findings of this study can help speech therapists and parents develop more understanding of the children's language development process.

**KEYWORDS:** Acquisition; Wh-words; Persian grammar; English grammar

## 1. Introduction

The way that a child acquires a language has always been a fascinating topic for research in different fields of study, particularly enjoying a specific focus on linguistics. Concerning the Persian language, a number of textbooks on Persian grammar have been published (e.g. Bateni 1995, Gholam-Alizadeh 1995, Lazard 1992, Mahootian 1997, Vahidian Kamyar & Emrani 2000) or the description of the Persian grammar has been provided through the analysis of adult Persian speakers (e.g. Dabir Moghaddam 1982, 1985; Rezaie & Zareifard 2016, Shamsavari & Dabir Moghaddam, 2016). In addition, an abundance of research has addressed the effects of Persian, as the first language, on the acquisition of a second or third language (e.g. Heidari Darani, 2012; Jabbari, 2018; Khany & Bazayr, 2013; Mollaie, et al., 2016; Youhannaie & Shoushtari, 1999). However, little is known about first language acquisition in Persian. To fill this gap, the present study aimed to provide an analysis of first language development in Persian, with a specific focus on the Wh-question formation.

As with other grammatical aspects, literature on Wh-structure involves providing a description of the structure in Persian (e.g. Kashefi, 2014; Karimi & Taleghani, 2007; Sadat Tehrani, 2011; Shiamizadeh, et al., 2017, 2018; Toosarvandani, 2008), a comparison of the structure in English and Persian (e.g. Abedi, et al., 2012; Gorjani, et al., 2012; Heidari Darani, 2015), or research on the effect of L1 on the acquisition of the structure in L2/L3 (e.g. Galbat & Maleki, 2014; Jabbari, et al., 2018; Mollaie, et al., 2016). However, research on how children acquire this structure is lacking. Hence, this paper aimed to

provide an analysis of the acquisition of the Wh-question formation and how variant it can be especially in the early stages. To provide more insightful information, a cross-linguistic comparison is also made with the process of Wh-structure acquisition in English.

## 2. Literature Review

### 2.1. A syntactic analysis of the structure of Wh-questions in English

According to Chomsky's minimalist program (Chomsky, 1995), questions are formed as a result of the syntactic presence in C-position of a strong question affix. According to Chomsky, Q is a strong affix, which requires checking via phonetic form. Accordingly, Q demands an overt head, which is realized by Subject- Auxiliary Inversion. In addition, Q carries a strong [+wh] specifier feature which must be checked by specifier-head agreement, forcing the movement of a Wh-operator to spec-CP. Hence, in English, object and most adjunct Wh-question formations go through three grammatical transformations. Initially, the object or adjunct of the declarative sentence (e.g. *Mary will buy the car*) is replaced by the Wh-word (e.g. *Mary will buy what?*). Next, the Wh-word moves to the beginning of the sentence (e.g. *what Mary will buy?*). Finally, there is the subject (*Mary*) and the auxiliary (*will*) inversion to produce the well-formed Wh-question (e.g. *what will Mary buy?*). In the case of subject Wh-questions, there is only the requirement for Wh-fronting and the subject-replacement by the Wh-word (e.g. *who bought the car?*).

### 2.2. A syntactic analysis of the structure of Wh-questions in Persian

Modern Persian, which belongs to the Indo-European family, is recognized as a pro-drop language. In Persian, pronominal subject omission is prevalent, and mostly subject is realized in the form of a verb suffix (Vaez Dalili, 2009). As an example, see the following sentence:

(1)

Ketabha ra avardand.

Books (Object marker) bring(Past).

They brought the books.

In this sentence, the subject (*anha* = they) is omitted, however it is easily identified by the native speakers from the verb ending "*and*". This pro-drop feature is also true about the interrogatives. That is, when constructing a Wh- or yes/no question, the subject can be dropped and the verb can act as the subject indicator, too.

Persian is generally categorized as a Wh-in-situ language (Lazard, 1992; Mahootian 1997; Youhanaei & Shoushtari, 1999). In Wh-in-situ languages, the Wh-element remains where it is produced in the deep structure. See the example (2) below:

(2)

Ali be madrese raft.

Ali to school went.

Ali koja raft?

Ali where went?

However, the state of Wh-structure in Persian being recognized as a Wh-in-situ language has been controversial. Some linguists, such as Kahnemuyipour (2001) and Adli (2010), believed that Farsi should not be considered solely as an in-situ language in the case of question formation. Persian permits two possibilities for some Wh-structures, and in some cases, it must move to the front position of the structure (immediately after or before the subject). See the example below:

(3)

Ali be madrese naraft chon mariz bud.

Ali to school did not go because he was ill.

Ali be madrese naraft, chera? \*

Ali to school did not go, why? \*

Deletion of the clause (*chon mariz bud*) and its replacement with the corresponding Wh-expression (*chera*) will result in an incorrect structure. In this example, the Persian grammar does not allow the Wh-word (*chera*) to remain in situ and triggers movement to the front or immediately after the subject.

(4)

*Chera Ali be madrese naraft?* (Why Ali to school did not go?)

*Ali chera be madrese naraft?* (Ali why to school did not go?)

Thus, in Farsi, it seems that Wh-movement to the front of the structure or after the subject is necessary in some cases. For Kahnemuyipour (2001), Persian cannot be considered either a language with syntactic Wh-movement or a Wh-in-situ language; rather, it is a language in which Wh-structure needs to undergo focus movement. According to Kahnemuyipour, in Persian, post-verbal Wh-arguments and Wh-adjuncts need to move to a pre-verbal position. Kahnemuyipour held that as [+wh] feature is weak in Persian, it must be triggered by an optionally realized strong [+focus] feature. As such, in Persian, the strong [+focus] feature is realized in Foc position, which is generated below CP and above TP. Consistent with this claim, Adli's (2010) study suggested that although Persian is claimed to possess a Wh-in-situ formulation, it presents a variety of word orders.

To add more into research on the formulation of the Wh-structure in Persian, the present study focused on how children acquire the structure. Other intricacies of Wh-question formation were also focused, and a comparison was made with the acquisition of the structure in English. In fact, the research aimed to set out the picture of an English and a Persian child's language development at approximately equal Mean Length of Utterance (MLU) values in the early stages (Brown, 1973), with a specific focus on detecting the patterns of correct use and errors in Wh-question acquisition, the type and frequency of the Wh-expression at each stage, and the sequence of the emergence of the variants across the two languages. More specifically, it was decided to:

1. What are the correct patterns of Wh-question structure in the production data of the? Persian and English children in general and at the specified developmental stages?
2. What is the frequency of Wh-words/variants at each specified stage and in what sequence do they emerge?
3. What are the two children's patterns of non-adult-like production, produced in their process of the acquisition of the WH structures?

### 3. Methodology

The data for this research came from the CHILDES database (<http://childes.talkbank.org>). The Persian data came from the Family Corpus, Lilia file. Lilia was a girl living in Tehran who was recorded at ages 1.11 to 2.10. The recordings were based on her mother and caretaker's interactions with Lilia, mostly in playtime activities at home. The English data were chosen from the English UK Corpus, Lara file. Lara was a girl living in England who was recorded at ages 1.09 to 3.03. Her grandmother and mother's interactions with Lara in their everyday activities were focused.

We used MLU (Mean Length of Utterance) as a simple way of making the two children's language production data comparable. In first language studies, MLU is a metric used to analyze and measure the complexity and development of a child's language skills. It provides valuable insights into how children acquire and progress in their native language. MLU is calculated by dividing the total number of morphemes (the smallest meaningful units of language) by the total number of utterances produced by the child. Morphemes can include words, prefixes, suffixes, and other grammatical elements. Utterances refer to complete statements or phrases made by the child. MLU is often used to assess various aspects of language development, such as vocabulary growth, grammatical complexity, and syntactic structures.

By analyzing MLU, researchers can track a child's language development over time. As children acquire language skills, their MLU tends to increase. This increase reflects their ability to produce longer and more complex utterances. MLU is also a valuable tool in language studies as it provides a quantitative measure to track and compare language development across different children and age groups. It helps researchers gain insights into the typical patterns of language acquisition, as well as identify any deviations or challenges that may require further attention or intervention.

In the present study, MLUs were calculated using the MLU function of the CLAN program (MacWhinney, 2000). The transcripts with MLU between 2 - 3.80 were analyzed for the purpose of the study. The data were categorized into stages based on Brown's (1973) criteria as it is presented in Tables 1 and 2.

**Table 1.** Brown's (1973) Classification of the stages for language development

Stages	Stage I	Stage II	Stage III	Stage IV
MLU	1.00-1.99	2.00-2.49	2.50-2.99	>3

**Table 2:** Classification of the research data into stages based on Brown's criteria

Child	MLU range (Stage II)	MLU range (Stage III)	MLU range (Stage IV)
<b>Lilia</b> (Persian)	2.020-2.460	2.60-2.830	3.520-3.820
<b>Lara</b> (English)	2.008-2.035	2.53-2.74	3.26-3.81

The data accessible were consistent with the second, third and fourth stages of Brown's criteria.

Repetitions, imitations, and incomprehensible structures were excluded from the analysis. The various patterns/variants were identified using content analysis. To address the reliability of coding and categorizations, inter-coder reliability (the author and a colleague) was performed. An initial analysis ensured a substantial kappa coefficient of agreement ( $k = .91$ ). Double checking, further discussions between the two coders, and expert consultation were employed to come to an agreement on the points of discrepancy.

## 4. Results and discussion

### 4.1. Patterns of correct use

The first research question aimed to identify the patterns of correct use in the Wh-question acquisition of both children. Regarding the Persian data, for the questions to be correct, the Wh-word needed to be in preverbal position. The differences could be in the linear order of the direct object and the Wh-word and the position of subject, which can be either preverbal or post-verbal. As Persian, to some extent, follows a free order, various patterns of Wh-structures were predicted to be produced by the child. To categorize the Wh-question variants in Persian, the position of the Wh-word, the subject, and the object were focused upon. Taking these positions into consideration, the Wh-patterns were classified into *wh-in-situ* and *wh-word-fronted*. In this categorization, the single-word Wh-question variants –e.g., *chi* (what) or *koja* (where)– and the single-word Wh-questions preceded by a proposition –as in *Az koja?* (From where?) – or by an adverb –as in *Pas chi?* (So what?)– were categorized as Wh-fronted. According to Toosarvandian (2008), the single Wh-word question, also referred to as "sluicing", is "an elliptical construction in which all of a constituent question goes missing except for the interrogative phrase" (P. 677). This movement-plus-deletion process is nearly identical in English and Persian; however, according to Toosarvandian (2008, p. 677), "in Farsi sluicing escapes deletion not by Wh-movement as in English but by a type of focus movement."

For the English data, in the case of the copula questions, the choice and placement of the Wh-word, copula, and subject needed to be correct. For the structures including auxiliaries, the accuracy of the choice and placement of Wh-word, auxiliary, main verb, and subject was required. Hence, compared to the Persian data, due to more constraints on word order in the Wh-structure, fewer patterns were predicted for the English data. The predictions came out to be true. Different patterns and their associated frequencies/percentages for both Persian and English variants are presented in Tables 3 and 4.

**Table 3.** Frequency and Percentage of the Wh- variants for the Persian data

<b>Wh-fronted</b>			
Variant	Examples	frequency	Percentage%
Wh (single word)	Chi? (what?)	44	33.08
PrepW/AdvWh	Az koja? (from where?) / pas chi? (therefore, what?)	6	4.51
Wh~V(single word, v=be)	Kodume?(which one is?)	8	6.01
WhV	Chi shod? (what happened?)	23	17.29
WhSV	Chera in mikhore?(why this fits?)	2	1.50
Wh~VS	Kie esmesh?* (what is name-EZ her)	3	2.25
WhS~V	Kodum ahang? (which song is?)	2	1.50
WhVO	Ki dad behesh?(who gave to her?)	3	2.25
Wh(OM)V	Chera ina ro nemibare? (why these(OM) does not take?*)	2	1.50
WhO(OM*)V	Chera inaro mizari?(why these put(2 <sup>nd</sup> SIG)?)*	7	5.26
WhSO(OM)V	Chera khale niloufar mobilesho var nadasht?( Why aunt Niloufar mobile-her did not take?)	2	1.50
Embedded wh	Bezar bebinam ke chi shode? (let see-1 <sup>st</sup> SIG what has happened?)	11	8.27
Total(Wh-fronted)		111	83.45
<b>WH-in-Situ</b>			
Variant	Examples	frequency	Percentage
SWh	To chi? (you what?)	4	3.00
SWh~V	In kie? (this who is?)	10	7.51
SWhV	To chi mikhori? (you what eat?)	2	1.50
AuxWhV	Mikhstim chi bokhorim?(want what eat?)	4	3.00
O (OM) Wh	Mami ro chi? (Mami (OM) what?)	1	0.75
Total( Wh-in-situ)		21	15.78
Total Wh-question		133	100

**Table 4.** Frequency and percentage of the Wh-variants for the English data

<b>Wh-fronted (Ex-situ)</b>				
Variant	Example	frequency	Percentage(%)	
Wh(single word)	Who?/why?	11	12.22	
WhAuxS	What's that? Where's daddy?	45	50.00	
WhAuxSV	Where does she go?	17	18.88	
WhPP	What about this one?	5	5.55	
embedded clause	SV(Wh+to+V)	I don't know how to make...	6	6.66
	SV(WhSV)	You know why it is really hard?	4	4.44
<b>Wh-Fronted (In-situ)</b>				
WhV	Who wants that?	2	2.22	
Total		90	100	

As the tables indicate, both the English and the Persian child used various types of Wh-questions; nevertheless, the Persian child seemed to produce more variable language than the English child. The single Wh-word and WhV were the most frequently used variants by the Persian child. Data also indicated the Persian<sup>1</sup> child's remarkably higher disposition to use Wh-fronted (83.45%), compared with the Wh-in-situ pattern (15.78 %). This seems contrary to the use of the structure by adult

<sup>1</sup> \* EZ stands for EZAFE; OM stands for Object Maker; SIG stands for Singular

language users, which is more frequently in the form of the Wh-in-situ structure. As such, it might be concluded that in the early stages of language acquisition, Wh-words are mostly produced in initial position.

The WH-question production of the English child, all Wh-fronted, were categorized into the *Wh-in-situ* questions –in the case of subject question– and *Wh-ex-situ* –in the case of object and adjunct questions. Wh-in-situ questions in English can also be produced by having the Wh-word in post-verbal position as in “you did what?”; however, no instances of such variant were observed in the production data of the English child studied in this research. The WhAuxV pattern, in which auxiliary is the copula *be*, seemed to be the most repeated variant produced by the child.

#### 4.2. Frequency and emergence order of different variants at each stage

Two issues were of interest in analyzing the data at each specific stage of the children's language development. The first was the issue of determining the dominant pattern/s of Wh-structure at each stage. The second was whether Wh-words/patterns are acquired in a particular order. Results for the frequency of each pattern are presented in Tables 5 and 6.

**Table 5.** Frequency of the Wh-variants at different stages for the Persian data

Variant	Wh(single)	prep/Adv/Wh	Wh~V	WhV	WhSV	Wh~VS	WhS~V	WhVO	WhSOV	Wh(OM)V	WhO(OM)V	Embedded wh	SWh	SWh~V	SWhV	AuxWhV	
Stage II	13	2	3	8	0	0	0	0	0	1	2	2	2	4	0	0	0
Stage III	13	1	5	10	0	0	0	2	0	0	3	3	1	3	0	2	1
Stage IV	18	3	8	5	2	3	2	1	2	1	2	6	1	3	2	2	0

**Table 6.** Frequency of the Wh-variants at different stages for the English data

Variant	Wh(single word)	WhAuxS	WhAuxSV	WhPP	WhV	Embedded Wh	
						Wh+to+V	WhSV
Stage II	2	32	0	0	0	0	0
Stage III	1	8	5	3	0	2	0
Stage IV	8	5	12	2	2	4	4

In the case of both children, the Wh-question variations increased steadily in number along the data points. In the case of the Persian child, single Wh-word and WhV were found to be the most frequently-used variants at all stages investigated. WhV pattern in Persian is possible for both object and subject questions; however, the pattern in English is only used to produce subject question. Single Wh-word question enjoyed a constant frequency of occurrence in the production data of the Persian child at different stages with the single Wh-word “*chi*” (what) frequently used at all stages and “*chera*” (why) at the third and fourth stages. There seemed to be an increase in the number and complexities of the embedded clauses at later stages, as well.

In the case of English, the second stage was predominantly characterized by the WhAuxS pattern, though including just a few fixed expressions (e.g. *what's that, where is...?*). This is in line with the Persian child's production at this stage. This can be explained via the pragmatic function of the Wh structure, which seems to be mostly aimed at obtaining information at earlier ages. In stage III, she seemed to produce a variety of questions but she occasionally omitted auxiliaries and/or did not have subject-auxiliary inversion. This stage was mostly characterized by the use of *be*, *do*, and *does* auxiliaries. By stage IV, more auxiliaries appeared and longer and more complex structures –including more embedded clauses were produced. As Table 6 indicates, embedded clauses at the third stage were mostly in the form of SV (how +to +V) as in “*you know how to make it*”. At stage IV, we also see instances of SV (Wh SV) structure, e.g., “*you know why it's really hard*”.

Concerning the second issue, i.e., the sequence of the emergence of the Wh-words (see Table 7), the Persian Child seemed to produce *chi* (what), *koja* (where), and *kodum* (which one) from the beginning of the data available. *Chera* emerged at the end of the second stage but seemed to be, along with *chi*, the most frequently used Wh-word in the following stages (i.e., stages III & IV). *Ki* (who) was used in all stages, but rarely. Other Wh-words were produced very rarely or not at all.

**Table 7.** Frequency of the use of Wh-words at each stage for the Persain data

Wh-word	Chi (what)	Chera (why)	Koja (where)	Ki (who)	Ku (where is)	Kodum (which one)
Stage II	28	2	2	3	0	3
Stage III	17	12	4	0	1	1
Stage IV	13	8	8	2	0	9

For the English child, in stage II, only *what* and *where* were used. *Why*, *How*, *When* and *Who* occurred very rarely in stage III, though, with the exception of *Who*, the other ones were more frequently used in stage IV (see table 8).

**Table 8:** Frequency of the use of Wh-words at each stage for the English data

Wh-word	what	why	where	who	how	when
Stage II	27	0	7	0	0	0
Stage III	7	1	5	1	2	2
Stage IV	3	8	14	1	6	3

### 4.3. Patterns of non-adult-like production

Another issue of interest in this study was to analyze the non-adult-like Wh-question production. Almost all such violations committed by the Persian child seemed to be related to the choice of the appropriate Wh-word. Examples are provided below:

(5)

a. kie\* esmesh?

Who is name her?

b. mami , mæge nemiduni emruz [/] emruz koja\* (*be jaye chandom*) Shæhrivære ?

Mami, don't you know today where (instead of what day of the month) Shahrivar is?

As mentioned earlier, Persian does not allow the placement of the Wh-word in post-verbal position. No instances of such error and other Wh-related syntactic errors were found in the data.

In the case of the English child, both lexical (Wh-word choice) and syntactic violations were observed

#### 4.3.1. Wh-word Choice Errors.

Most of the lexical violations were pertinent to the misapplication of the Wh-word *What* instead of *Who*. See examples *a*, *b*, and *c*, below:

(6)

a.

CHI: what\*'s that?

MOT: who is it?

MOT: It's Ronald McDonald.

*b.*

CHI: what\*'s that?

MOT: who is it?

CHI: what\*'s that ?

MOT: a little girl.

*c.*

CHI: what\*'s that ?

MOT: what's what?

CHI: that boy

As mentioned in the previous section, *Who* emerged later in the child's acquisition process of Wh-words and even in later stages, it was used rarely. An interesting point is that in stage III, the child did the other way around and misplaced *what* with *who*. See example *d*, below:

(7)

CHI: who\* are you doing?

I'm doing Cesca. MOT:

This might indicate that the child is having challenges with the structure.

### 4.3.2. Syntactic Errors.

The following types of syntactic error were found in the production data of the English child.

#### 4.3.2.1. Auxiliary omission

Auxiliary omissions refer to the violations in which the auxiliary is dropped. In such cases, tense is not overtly marked on the main verb. See the examples in the production data of Lara.

(8)

*a.*

CHI: when after finished we go?

MOT: pardon?

CHI: when [/] when after finished this apple I go?

MOT: go where?

CHI: go and get daddy

*b.*

MOT: my poor sheep.

CHI: where [?] mummy's poorly sheep gone?

MOT: you've destroyed it.

*c.*

MOT: you want to do this page?

MOT: right.

CHI: where it gone?



d.

CHI: who drink my tea?

CHI: who drink my tea ?

#### 4.3.2.2. *Inversion errors*

(no subject and auxiliary/copula inversion). For example:

(9)

MOT: oh no.

MOT: so then they have some dinner.

CHI: what [//] why he hasn't got any dinner?

MOT: he's not hungry (be)cause he's sad.

#### 4.3.2.3. *Agreement errors*

For example:

(10)

a.

CHI: who want [\*] to go up the ladder?

CHI: let me make bigger than...

CHI: that's bigger than sky.

b.

CHI: where's [\*] lights?

CHI: know where the lights is?

CHI: lights.

Dad: we've gotta ask mummy

There were also instances of violation for which a particular pattern could not be specified. See the example below:

(11)

MOT: they're going to bed.

CHI: who?

MOT: monkey's going to bed.

CHI: what he's [\*] bed?

Taken together, based on the results obtained from the analysis of the two children's non-adult-like production, it can be concluded that in the case of Persian, the child was able to produce correct (adult-like) questions from the very beginning of the multi-word speech stage. However, this was not equally true of the English child who seemed to produce more syntactic errors even at the later stages of her language development. This is not surprising as forming English Wh-questions are more demanding as they should go through obligatory movements and require inversion, agreement, and auxiliary checking. Patterns of such violations in the English data seemed to agree that the acquisition of WhAuxS (the most frequently used pattern from

the first stage) was the easiest acquired Wh-structures. Besides, the data indicated only sporadic instances of inversion errors in their production data, that is, the child produced both un-inverted and well-formed structures throughout the same period. Such errors were observed even at later stages meaning that there was no evidence of a specific un-inversion stage. The same was true for auxiliary omission errors.

## 5. Conclusion

This research aimed to provide a picture of the acquisition process of the Wh-structure in Persian and English. Through the study, we compared the data production of the two children in terms of type and frequency of the correct patterns in general and at comparable stages. Non-adult-like forms of Wh-expressions were also analyzed. Different patterns for both languages were found; however, the data suggested Persian to be more variant compared to English. Another interesting finding was that unlike most claims in literature concerning the Wh-in-situ status of Wh questions in Persian, the present study indicated it to be more Wh-fronted in the case of the studied child. Preferable patterns and Wh-words at each stage were compared. The data also suggested that the Persian child had an easier path to acquire the structure compared with the English child who seemed to have more challenges and produced more syntactic violations. Findings can contribute to the existing body of knowledge on English/ Persian language acquisition and bilingualism, as well. Besides, this study can help speech therapists and parents develop more understanding of the children's language development process. However, caution should be taken in making generalization about children's acquisition of Wh-question structure for two reasons: first, the sampled data analyzed in this study comprises just a small portion of the child's speech and second, the characteristics observed might be merely individual. Hence, future studies employing larger data sets at various developmental stages and comparing sample data from more children are suggested.


## 6. References

- Abedi, F., Moinzadeh, A., & Gharaei, Z. (2012). WH-movement in English and Persian within the framework of government and binding theory. *International Journal of Linguistics* 4, 419–432. <https://doi.org/10.5296/ijl.v4i3.2325>
- Adli, A. (2010). Constraint cumulativity and gradience: Wh-scrambling in Persian. *Lingua* 120, 2259-2294. <https://doi.org/10.1016/j.lingua.2010.01.003>
- Batani, M. R. (1995). *Tosif-e Sakhteman-e Dastury-e Zaban-e Farsi [Description of Persian Syntax]*. Amir Kabir Publishers.
- Brown, R. (1973). *A first language: The early stages*. Harvard University Press.
- Chomsky, N. (1995). *The minimalist program*. MIT Press.
- Dabir-Moghaddam, M. (1985). The Persian Passive. *Journal of Linguistics* 3, 31-46.
- Dabir-Moghaddam, M. (1982). *Syntax and Semantics of Causative Constructions in Persian*. [Dissertation]. [Urbana-Champaign]: University of Illinois.
- Galbat, M., & Maleki, R. (2014). Acquisition of Wh-questions in English by Persian Monolinguals and Iranian Arab Bilinguals. *Theory and Practice in Language Studies* 4.2, 1421-1433. <https://doi.org/10.4304/tpls.4.7.1421-1433>
- Gholam-Alizadeh, Kh. (1995). *The Structure of Persian*. Ehyaye Ketab.
- Gorjian, B., Naghizadeh, M., & Shahramiri, P. (2012). Making interrogative sentences in English and Persian language: A contrastive analysis approach. *Journal of Comparative Linguistics and Literature* 2, 120–124.
- Heidari-Darani, L. (2015). Language and Variation: A Study of English and Persian Wh-questions. *Journal of Applied Linguistics* 8.16, 37-66.
- Heidari-Darani, L. (2012). Persian-English Interlanguage Wh-questions: Do They Experience Patterned Variation? *Archives des Sciences* 65.7, 28-41.
- Jabbari, A. A. (2018). Acquisition of Noun Modifiers in Third Language Arabic (L3) by Iranian Persian (L1) Speakers Learners of English as a Second Language. *Foreign Language Research Journal* 8.1, 83-104.

- Jabbari, A. A., Achard-Bayle, G., & Ablali, D. (2018). Acquisition of L3 French wh-question structure by Persian-English bilinguals, *Cogent Education* 5.1, 1524551, <https://doi.org/10.1080/2331186X.2018.1524551>
- Kahnemuyipour, A. (2001). On wh-questions in Persian. *Canadian Journal of Linguistics* 46.1/2, 41-61. <https://dx.doi.org/10.1017/S000841310001793X>
- Karimi, S., & Taleghani, A. (2007). “Wh-movement, interpretation, and optionality in Persian,” in Phrasal and clausal architecture: Syntactic derivation and interpretation, eds S. Karimi, V. Samiiian, & W.K. Wilkins, 167-187. Benjamins.
- Kashefi, G. (2014). An Integrated Minimalist Analysis on the Position of Colloquial Persian Wh- phrases. *International Journal of Literature and Art (IJLA)* 2, 25-31.
- Khany, R., & Bazayr, A. (2013). A Generative Analysis of the Acquisition of Negation by Iranian EFL Learners: A Typological Study *RALS* 4.1, 62-87.
- Lazard, G. (1992). *A Grammar of Contemporary Persian*. English translation. Costa Mesa, California: Mazda. (Translated from French by Shirley Lyons; first published in 1957 as *Grammaire du persan contemporain*). Klinksieck.
- Mahootian, SH. (1997). *Persian descriptive grammars*. New York: Routledge.
- Mac Whinney, B. (2000). *The CHILDES Project: tools for analyzing talk*, 3rd ed. Lawrence Erlbaum Associates.
- Mollaie, A., Jabbari, A. A., & Rezaei, M.J. (2016). The acquisition of French (L3) Wh-question by Persian (L1) Learners of English (L2) as a Foreign Language Optimality Theory. *International Journal of English Linguistics* 6.7, 96-47. <https://doi.org/10.5539/ijel.v6n7p36>
- Rezaie, M. & Zareifard, R. (2016). The structure of Persian simple sentences. *Journal of Educational and Social Research* 6(10), 94-101. <https://doi.org/10.5901/jesr.2016.v6n1p94>
- Sadat Tehrani, N. (2011). The intonation patterns of interrogatives in Persian. *Linguistic Discovery Journal* 9, 105–136. <https://doi.org/10.1349/PS1.1537-0852.A.389>
- Shahsavari, A., & Dabir-Moghaddam. (2016). Transitivity in Persian based on the Minimalist Approach. *Persian Language and Iranian Dialects* 1.1, 139-149. <https://dorl.net/dor/20.1001.1.65852476.1395.1.1.8.6>
- Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2017). The role of prosody in identification of Persian sentence types: Declarative or wh-question? *Linguistics Vanguard*, 2016- 2085. <https://doi.org/10.1515/lingvan-2016-0085>
- Shiamizadeh, Z., Caspers, J., & Schiller, N. O. (2018). Do Persian native speakers prosodically mark wh-in-situ questions? *Language and Speech* 62.2, 231-249. <https://doi.org/10.1177%2F0023830917753237>
- Toosarvandian, M. (2008). Wh-movement and the syntax of sluicing. *Journal of Linguistics* 44, 677-722. <http://doi.org/10.1017/S0022226708005367>
- Vaez Dalili, M. (2009). Agreement (AGR) and the pro-drop/non-pro-drop variation: A meta-analysis of GB and MP accounts. *PhiN* 49, 84-105.
- Vahidian Kamyar, T., & Emrani, G.R. (2000). *Persian Grammar*. SAMT.
- Youhanaie, M., & Gouniband Shoushtari, Z. (1999). Farâgirie sâxtâ r hâ ye porseshie sâdeyebabâne Engelisi tavasote gooyeshvarâne Fârsi zabân va dozabâne hâye Arab va Fârsi [Acquiring simple English interrogative structures by Persian speakers and Arab-Persian bilinguals]. *Elmi- pazhooheshi Journal of Department of Literature and Humanities* 2.52, 125-137.



## The Effect of Online Quizlet Flashcards and Student-Created Quizlet Flashcards on Learning and Retention of Productive and Receptive Vocabulary Knowledge of Elementary Iranian EFL Learners

Hossein Akbari<sup>1\*</sup>   
Farrokhlagha Heidari<sup>2</sup>

\*Ph.D. Candidate in Applied Linguistics, Department of English Language and Literature, University of Yazd, Yazd, Iran

<sup>2</sup>Assistant Professor in Applied Linguistics, Department of English Language, University of Sistan and Baluchestan, Sistan and Baluchestan, Iran

### ARTICLE HISTORY

Received: 16 April 2023  
Revised: 10 May 2023  
Accepted: 30 August 2023  
Published: 30 September 2023

### CORRESPONDING AUTHOR

E-mail: [akbarri.hossein@gmail.com](mailto:akbarri.hossein@gmail.com)

### ABSTRACT

Learning vocabulary seems to be a major focus of language instruction. The role of vocabulary knowledge in language learning has been studied by many researchers. The present study was conducted to investigate the effect of using online Quizlet flashcards and student-created Quizlet flashcards on learning and retention of receptive and productive vocabulary knowledge. Sixty Iranian elementary high school students who were randomly divided into 3 groups (twenty subjects per group) participated in this study. The materials and data collection instruments included the Quick Placement Test (QPT), the students' textbook, Top Notch1 from which the vocabulary items were selected, a pre-test, an immediate post-test, and a delayed post-test. It should be mentioned that the control group was taught the words through the conventional method while the first experimental group received the words using online flashcards and the next experimental group worked on developing their own Quizlet flashcards. The results showed that in terms of receptive vocabulary knowledge learning, the students of the online Quizlet flashcards group significantly outperformed students of the control group and online Quizlet flashcards. In learning of the productive vocabulary knowledge, the learners of the control group performed significantly better than other two experimental groups. The results on vocabulary retention revealed that students of the student-created Quizlet flashcards group significantly outperformed students of the control group in terms of retention of the receptive vocabulary knowledge, and in terms of the productive vocabulary knowledge retention there were no significant difference between the groups of the study. These findings will be advantageous for English language teachers to choose the teaching techniques which can facilitate the process of language teaching and learning.

**KEYWORDS:** Online flashcards; Student-created flashcards; Receptive vocabulary; Productive vocabulary; Vocabulary learning; Vocabulary retention

## 1. Introduction

Mastery of target language necessitates learning many aspects of that language including grammar, pronunciation, writing, pragmatics, and so on, but there is no doubt that vocabulary is the most important aspect of a language that learners need to learn (Folse & Briggs, 2004). For the purpose of second or foreign language learning and the ability of communicating

successfully, one needs to learn the skills of listening, writing, reading, and speaking. The role of vocabulary is indisputable in developing four main skills (Farrokhi et al., 2021).

The advent of new technologies has brought many changes in the field of English teaching and English learning. The traditional methods are not adequate in the term of language learning, so applying new methods is required to enhance learners' motivations. The role of CALL in vocabulary learning is crucial since it provides students with metacognitive activities allowing for making inferences through contexts using imagery and semantic techniques (Köse et al., 2016).

As Schmitt (2010) states, an important way to convey new vocabulary is the use of learners' independent strategies. CALL materials encourage students to develop their independency, help them to find their needs in language learning (Lee, 2004). When teachers use CALL in the process of their teaching, learners get motivated highly and consider CALL as a profitable in term of language learning (Afshari et al., 2009). Many teachers look at students' independency as a threat to their authority in the class. They are more satisfied to hold all procedures in their hands.

The use of CALL among Iranian instructors is not an option. Unfortunately, a large portion of students and teachers are not trained enough to use these new technologies. Due to the fact that CALL is a newborn field in language teaching, comparing the traditional methods of teaching with these newly developed methods need more consideration. Besides, most of the previous papers and articles directed their attention toward teacher-made materials. Little attention has been given to the student-generated materials in CALL. Therefore, the present study tried to accomplish a twofold purpose; first, it was a general comparison between conventional methods of vocabulary learning. Second, this study tried to compare student-generated materials and teacher-developed materials in the field of CALL.

## 2. Literature review

### 2.1. Theoretical framework

#### 2.1.1. Flashcards

In order to discuss a term first we need to define it. Based on Hung (2015, p. 1) "Word cards are a set of double-sided cards designed for direct learning of vocabulary that allow learners to practice form-to-meaning and meaning-to-form recall in repeated retrieval of L2 words, by flipping the front and back sides of the cards."

Nicholson (1998) states that flashcards can be joyful and funny for learners in the process of learning. Flashcards can also have positive effects on the visual learning of learners because they are clear and colorful. The flashcards are used not only for teaching vocabulary meaning but also for teaching prepositions, articles, sentences structure, and phrasal verbs (Tan & Nicholson, 1997).

Assigned flashcards learning into decontextualized learning category. They also regard that despite their separation from context, the flashcards are popular among language learners, especially for those learners who needs a material for vocabulary self-testing.

While flashcards had been used by learners for a long time, recent advances in technology have lead learners and teachers to take advantage of using digital or online flashcards. Using flashcard can benefit the learners in two ways: First, the studies showed intentional learning takes less time in comparison to incidental learning which needs long-time exposure (Hung, 2015). Second, it can encourage learners to be autonomous in their learning.

#### 2.1.2. Receptive and productive vocabulary knowledge

Laufer (1998) determines that "the process of learning a second language has often been described as the learner's progress along the interlanguage continuum from a non-existent knowledge towards native-like competence without necessarily reaching it" (p.1). This suggests that language learning involves a gradual process. The development of vocabulary is not only a matter of expanding one's vocabulary size but also of deepening one's familiar words. Although the distinction between breadth of the words and depth of the words has been known, the studies on this area are still neglected (Verhallen & Schoonen, 1998).

Not all words that a learner hears in a lesson or other lessons become his active vocabulary knowledge. Some words in the process of vocabulary learning remain passive and learners will understand their meaning only when they hear them or read them, but they are not part of active vocabulary knowledge; that is, learners are not able to use them in their speaking or writing. The vocabulary should be presented and practiced systematically for active use (Finocchiaro, 1986).

The receptive/productive vocabulary knowledge distinction depends on the distinction between receptive skills of listening and reading, and productive skills of writing and speaking. The listening and reading skills are called receptive skills due to the fact that we receive input from others. Productive skills imply the idea that we produce language forms using writing and speaking to convey a message to others (Nation, 2001). This distinction is not completely suitable because there are some productions in receptive skills and vice versa.

### 2.1.3. Vocabulary retention

Vocabulary retention has been long neglected in the field of language learning. Richards and Schmidt (2002) define vocabulary retention as “the ability to recall or remember things after an interval of time. In language teaching, retention of what has been taught may depend on the quality of teaching, the interest of the learners, or the meaningfulness of the materials” (p. 457).

Retention is the process of transferring new information into long-term memory. This means you’ve effectively taken in the information and are able to recall it in the future (Folse, 2006). Retention needs to engage more in the process of learning and acquiring words through different channels and applying them in different contexts (Hummel, 2010). Vocabulary learning is strengthened when learners use visual and linguistic methods in learning a word. It will be also helpful to associate them with similar words (Tinkham, 1997).

## 2.2. Empirical studies

In a study related to using flashcards, Komachali and Khodareza (2012) investigated the effect of using flashcards on vocabulary learning of Iranian pre-university students. The participants of the study were 50 female students. They were randomly divided into two groups. The control group received traditional treatment and the experimental group received flashcard treatment. The scores of the post-test revealed that there was a significant difference between the two groups. The results showed that compared to the control group the experimental group had a higher level of improvement in their vocabulary knowledge.

Furthermore, Mohammadnejad et al. (2012) tried to study the efficiency of word lists and picture flashcards and their effects on students’ English vocabulary learning. There were 36 students in an English institute. The researchers used the two mentioned techniques to teach new words to students. The results of the post-test indicated that using flashcard was more helpful than word list.

In another study in line with these studies, Sitompul (2013) investigated the effect of using flashcards and word lists on vocabulary learning. The students of two classes were participants in this study. The experimental group received flashcard treatment and the control group received word lists treatment. The findings of the study showed that the vocabulary knowledge of students improved after using flashcards. It was found that using flashcards could help students to memorize the words more easily.

In line with previous studies, Mojarradi (2014) compared the effect of using flashcards on ESL students’ ability to learn vocabulary. The participants of the study were 40 pre-university students in two separate classes. The participants were given a post-test to make sure their vocabulary level was at the same level. The control group was asked to develop their own flashcards for learning the new vocabulary and the experimental group was asked to use the already prepared flashcards. The results of the post-test showed that the group with already prepared flashcards was able to learn vocabulary better than students in the control group who made their own flashcards. The findings of this study showed the positive effect of flashcards on vocabulary learning.

Osman et al. (2015) investigated the position of flashcards vs. word lists for vocabulary building. There were ninety students from King Abdullah High school. There were three experimental groups: 1. the students in the first group studied vocabulary from bilingual word lists. 2. The students of the second group used bilingual flashcards, where the target word and its Arabic meaning were on one side. 3. The last experimental group studied words on bilingual flashcards, where the target word was on one side, and the Arabic meaning was presented on the other side. The study followed the T1-treatment-T2 format, where participants were pretested with the target words, given the new vocabulary, and post-tested on the target words for any retention or attrition effects. Finally, the researchers concluded that there were no significant differences between the three types of flashcards in King Abdallah High School.

Taghizadeh and Porkar (2018) compared the effect of tablets, Short Message Service (SMS), and flashcards on learner’s vocabulary knowledge and their attitude. The participants were 45 Iranian advanced learners. The participants were divided into three groups. The first group studied the vocabulary via SMS, next group studied vocabulary using tablet and the last group learned vocabulary items via flashcard. The results of the pre-test showed that vocabulary instructions using SMS, tablet, and flashcard was effective in improving learners’ vocabulary knowledge. Most of the learners in the flashcard group considered flashcards as an effective and flexible tool for learning words.

Based on the results of some studies, some researchers disagreed with the use of flashcards due to the fact they considered flashcards as tools focused on memorization. Basoglu and Akdemir (2010) compared the effect of flashcard and a mobile program on learners' vocabulary development. The participants of this study were 60 university students. The results of the study showed that the mobile phone program was a more effective tool than flashcard in developing students' vocabulary.

In another study, Azabdaftari and Mozaheb (2012) studied the effect of mobile learning and flashcard on vocabulary knowledge. The participants of the study were 80 university students studying English literature and translation at the BA level. The participants were divided into two groups of forty. Forty students were in the experimental group learning vocabulary using mobile phones. The students of the control group learned the new words using flashcard. A multiple-choice test was used to assess after experiment. The result revealed that compared to flashcard using mobile phones would be more beneficial.

In addition, Khodashenas et al. (2014) compared the effect of using flashcards and educational cartoons on the vocabulary learning of intermediate EFL learners. There were 44 students in this study as participants. They were selected based on a pre-test and divided into a control group and an experimental group. The experimental group was presented words using Magic English cartoons while the control group was taught words using flashcards. The result of post-test indicated a significant difference between the performances of students in two groups. The students in experimental group outperformed the students of the control group. It was concluded that using magic cartoons could be seen as effective technique to improve learners' knowledge.

### 3. Methodology

#### 3.1. Instrumentation

The instruments used in this study are presented in below. They include a placement test, pre-test, post-test, and a delayed post-test.

##### 3.1.1.Placement test

In order to homogenize participants according to their level of language proficiency, a standard language proficiency test, i.e. quick placement test of Oxford University Press and University of Cambridge Local Examinations Syndicate was used. This test includes 60 multiple choice questions of vocabulary, grammar, and comprehension passage. All participants were given the same placement test to check if the learners' proficiency level is suitable for the present study. According to the test scale, the participants who achieved 21-40 were considered as elementary learners and were appropriate for the purpose of this study.

##### 3.1.2.Pre-test

The pre-test was researcher-made and aimed to compare the results with the post-test to determine whether there was any significant progression in learners' receptive and productive vocabulary knowledge. The test had two parts. In the first part there were 15 matching questions. The students were asked to match each word on the left column to its definition in the right column. This part was designed to measure learners' receptive knowledge of words. Next part includes 15 fill in the blanks questions and in each part the first three letters of the words were given. The second part was designed to measure learners' productive knowledge. The pre-test was designed based on the content of the course which was from book entitled *Top Notch 1 (Third edition)* by (Saslow & Ascher, 2011). In order to measure the reliability of the teacher-made pre-test, it was administrated to a class of 20 students who had similar characteristics and proficiency level with the learners in control group and experimental groups. The reliability of the pilot study calculated by Cronbach's alpha turned to be 0.82 which considered a reliable test.

##### 3.1.3.Post-test and delayed post-test

In this study, the pre-test, post-test, and delayed post-test were exactly the same. The purpose of conducting post-test was to measure the possible difference in receptive and productive knowledge of learners after receiving treatment. One of the major purposes of this study was measuring the effect of treatments on retention of receptive and productive vocabulary knowledge. For achieving this purpose, a delayed post-test was conducted after two weeks from the first post-test to measure the level of retention in learners

### 3.2. Materials

#### 3.2.1.Text book

Top Notch 1 was the book from which words were selected. This book consists of 10 units, from each lesson of this book, 15 words were selected that were more important based on the researcher's opinion and were in line with the main objectives of that lesson. The reason for choosing this book was the researcher's mastery over the topics of the book due to its teaching in the classrooms and also its compatibility with the level of students in this research. The units of the book include the following topics:

- Unit1: Getting Acquainted
- Unit2: Going out
- Unit 3: The extended family
- Unit 4: Food and Restaurants
- Unit 5: Technology and You
- Unit 6: Staying in Shape
- Unit7: On Vacation
- Unit8: Shopping for Clothes
- Unit9: Taking Transportation
- Unit10: Spending Money

### 3.2.2. Quizlet flashcards

In order to make and use online flashcards, the Quizlet website (<https://www.quizlet.com>) was used. The website is designed to enable learners to find suitable courses based on their needs or they can create their own courses. The website has different modes for learning and studying new words. The modes include:

- Flashcards: In this mode, users are shown a card for each term, which they can flip the cards. The face of the card can be an image, a word, or both.
- Learn: In this mode, learners answer multiple-choice and written questions repeatedly.
- Write: In this mode, Learners are given a definition or term and should write a definition or term that goes with what is shown.
- Spell: This mode enables the learners to work on spelling of the words. The term is read out loud and learners must type what they hear with correct spelling.

Using Quizlet website helps learners to work on different aspects of a word. They can easily receive the pronunciation, spelling, and meaning of the words and practice all these aspects in different modes of the website.

### 3.3. Data collection procedure

For conducting this study, three groups of learners were needed: The first group which considered as control group taking traditional method (explicit teaching of vocabulary, the second group which is the first experimental group were responsible to work on teacher-made flashcards, and the third group which was the second experimental group were asked to create their own online flashcards. The first step was conducting a placement test to measure students' proficiency level and select the students with appropriate knowledge for this study (Elementary learners).

At the first step the students should be familiar with this website and application and its features. After the participants' familiarization with both site and application, the pre-test will be administered to three groups. The pre-test and post- test for this study will be the same. In the first part of the test students should match 15 words to their definition and this part is designed to measure the students' receptive vocabulary knowledge. In the second part the students should fill in the blanks with some



words which will be given to them in the box. The second part of the test will consist of 15 sentences and this part will be for measuring students' productive vocabulary knowledge. For teaching the vocabulary, Top Notch1 (third edition) which has suitable level will be used. After the pre-test is done, the participants are supposed to elicit the meaning of the words of each unit from the context and will receive feedback from the teacher. The course will be conducted in 5 weeks and 2 sessions on every week. Each session 1 unit of the book will be covered. The first group of students which are experimental group1 are supposed to work on the flashcards which will be uploaded by the teacher on the Quizlet. The experimental group2 will develop their own flashcards on the Quizlet based on what they will be taught in the class and the third group which is our control group will work on the definitions and examples they will be provided in the class.

After 10 weeks the students of the three groups will have an immediate post-test. For measuring the retention of productive and receptive vocabulary knowledge, the posttest will be readministered 2 weeks after conducting the first post-test.

### 3.4. Data analysis procedure

The main focus of this study was to investigate the effect of online Quizlet flashcards and student-created Quizlet flashcards on learning and retention of receptive and productive vocabulary knowledge of Iranian EFL learners.

Concerning the first research hypothesis, the participants' scores in pre-test and post-test receptive vocabulary knowledge were drawn. Firstly, a One-way ANOVA was done on the scores of the pre-test receptive vocabulary knowledge to ensure that there was no significant difference in the scores of the participants prior to the intervention. In order to determine the effect of online Quizlet flashcards and student-created Quizlet flashcards on learning of receptive vocabulary knowledge, a One-way ANOVA was done on the post-test receptive vocabulary knowledge scores in the three groups. Afterwards, to determine which group had the most improvement in learning of the receptive vocabulary knowledge, LSD post-hoc test was conducted on the results of post-test receptive vocabulary knowledge.

For the second research hypothesis, similar to the first question, a One-way ANOVA was done on the results of the pre-test productive vocabulary knowledge to ensure homogeneity before the treatment. Afterward, a One-way ANOVA was conducted on the post-test productive vocabulary knowledge scores in three groups of learners. In order to show which group had the most improvement in productive vocabulary knowledge, LAD post-hoc test was used.

For the third research hypothesis, a One-way ANOVA was conducted on the delayed post-test receptive vocabulary knowledge scores in three groups of learners. In order to show which group had the most ability in retention of receptive vocabulary knowledge, LSD post-hoc test was employed.

For the fourth research hypothesis, similar to the third research hypothesis, a One-way ANOVA was conducted on the delayed post-test productive vocabulary knowledge scores in three groups of learners. In order to show which group had the most ability in retention of productive vocabulary knowledge, LSD post-hoc test was employed.

## 4. Results and discussion

### 4.1. Results

The main objective of this study was to investigate the effect of Quizlet flashcards on learning and retention of receptive and productive vocabulary knowledge of students. This section examines the hypotheses raised in this study to examine their acceptance or rejection.

#### 4.1.1. Research hypothesis 1

H01. There is no significant difference between traditional method, online Quizlet flashcards, and student-created Quizlet flashcards on learning of receptive vocabulary knowledge.

To evaluate the effectiveness of the treatment on the learning of receptive knowledge and to test the first hypothesis, another One-way ANOVA was run on participants' post-test receptive knowledge scores. The descriptive statistics of the post-test scores are shown in Table 1.

**Table 1.** Descriptive statistics of the post-test receptive vocabulary knowledge scores

Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control	20	10.2500	2.65320	.59327	9.0083	11.4917	6.00	15.00
Experimental 1	20	12.6000	1.66702	.37276	11.8198	13.3802	9.00	15.00
Experimental 2	20	10.6500	2.15883	.48273	9.6396	11.6604	8.00	15.00
Total	60	11.1667	2.39467	.30915	10.5481	11.7853	6.00	15.00

As it can be seen in Table 1, the mean of all three groups has increased in the post-test in terms of receptive knowledge. In order to investigate whether the increase in the receptive knowledge was significant or not, the researcher conducted a One-way ANOVA on the post-test scores of the receptive vocabulary knowledge. The results of the one-way ANOVA on the post-test receptive vocabulary knowledge scores are shown in Table 2.

**Table 2.** One-way ANOVA: comparison between scores on post-test receptive vocabulary knowledge scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	63.233	2	31.617	6.551	.003
Within Groups	275.100	57	4.826		
Total	338.333	59			

As the results of Table 2 show, an F ratio of  $F(2, 59) = 6.551$ ,  $p = .003 < .05$  indicated a statistically significant difference among the three groups; therefore, the first null hypothesis of the study was rejected. To find the location of the difference a post-hoc LSD was carried out. The results are presented in Table 3.

**Table 3.** LSD post-hoc test: comparison between post-test receptive vocabulary knowledge scores

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.
Control	Experimental 1	-2.35000*	.69472	.001
	Experimental 2	-.40000	.69472	.567
Experimental 1	Control	2.35000*	.69472	.001
	Experimental 2	1.95000*	.69472	.007
Experimental 2	Control	.40000	.69472	.567
	Experimental 1	-1.95000*	.69472	.007

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

The mean difference reported in Table 3. indicates a statistically significant difference between the online Quizlet flashcards (Experimental group 1) and control groups. There is also a significant difference between online Quizlet flashcards and student-created Quizlet flashcards (Experimental group 2) groups as presented in Table 4. the mean difference of  $M = 1.95$ ,  $p = .007 < .05$ . The most difference lies between the experimental group 1 and the control group with a mean difference of  $M = 2.35$ ,  $p = .001 < .05$ . Based on the mean differences, the three groups can be ordered in this way with regard to receptive knowledge scores: (1) online Quizlet flashcards (2) student-created Quizlet flashcards (3) control. Moreover, there is not a significant difference between the student-created Quizlet flashcards (Experimental group 2) and control group' scores on post-test receptive vocabulary knowledge as presented in Table 4. the mean difference of  $M = .40$ ,  $p = .567 > .05$ .

#### 4.1.2. Research hypothesis 2

To evaluate the effectiveness of the treatment on learning of productive vocabulary knowledge and to test the second hypothesis, another One-way ANOVA was run on participants' post-test scores on productive vocabulary knowledge. The descriptive statistics of the post-test scores are shown in Table 4.

**Table 4.** Descriptive statistics for the post-test productive vocabulary knowledge scores

Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control	20	7.7000	1.92217	.42981	6.8004	8.5996	5.00	13.00
Experimental 1	20	5.8500	2.96071	.66203	4.4643	7.2357	2.00	13.00
Experimental 2	20	6.3000	2.40832	.53852	5.1729	7.4271	3.00	13.00
Total	60	6.6167	2.55178	.32943	5.9575	7.2759	2.00	13.00

As it can be seen in Table 4., the mean of all three groups has increased in the post-test in terms of productive vocabulary knowledge. In order to investigate whether the increase in the productive vocabulary knowledge was significant or not, the researcher conducted a One-way ANOVA on the post-test scores of the productive vocabulary knowledge. The results of the one-way ANOVA on the post-test productive vocabulary knowledge scores are shown in Table 5.

**Table 5.** One-way ANOVA: comparison between scores on post-test productive vocabulary knowledge scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	37.233	2	18.617	3.059	.042
Within Groups	346.950	57	6.087		
Total	384.183	59			

As the results of Table 5. show, an F ratio of  $F(2, 59) = 3.059$ ,  $p = .042 < .05$  indicated a statistically significant difference among the three groups; therefore, the second null hypothesis of the study was rejected. To find the location of the difference a post-hoc LSD was carried out. The results are presented in Table 6.

**Table 6.** LSD post-hoc test: comparison between post-test productive vocabulary knowledge scores

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.
Control	Experimental 1	1.85000*	.78018	.021
	Experimental 2	1.40000	.78018	.078
Experimental 1	Control	-1.85000*	.78018	.021
	Experimental 2	-.45000	.78018	.566
Experimental 2	Control	-1.40000	.78018	.078
	Experimental 1	.45000	.78018	.566

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

The mean difference reported in Table 6. Indicates a statistically significant difference between the Control group and online Quizlet flashcards group, as presented in Table 4.6 the mean difference of  $M = 1.85$ ,  $P = .021 < .05$ . There is not a significant difference between control group and student-created Quizlet flashcards group as Experimental group 2, as presented in Table 6. the mean difference of  $M = 1.40$ ,  $P = .078 > .05$ . Also, there is not a significant difference between online Quizlet flashcards (Experimental group 1) and student-created Quizlet flashcards (Experimental group 2), as presented in Table 6, the mean difference of  $M = .45$ ,  $P = .566 > .05$ . Based on the mean differences, the three groups can be ordered in this way with regard to productive knowledge scores: (1) Control group (2) student-created Quizlet flashcards (3) online Quizlet flashcards.

### 4.1.3. Research hypothesis 3

H03. There is no significant difference between traditional method, Online Quizlet flashcards, and Student-Created Quizlet flashcards on retention of receptive vocabulary knowledge.

To evaluate the effectiveness of the treatment on retention of receptive vocabulary knowledge and to test the third hypothesis, One-way ANOVA was run on participants' delayed post-test scores on receptive vocabulary knowledge. The descriptive statistics of the post-test scores are shown in Table 7.

**Table 7.** Descriptive statistics for the delayed post-test receptive vocabulary knowledge scores

Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control	20	7.9500	2.66508	.59593	6.7027	9.1973	5.00	15.00
Experimental 1	20	11.3000	1.75019	.39135	10.4809	12.1191	9.00	15.00
Experimental 2	20	9.3000	3.14726	.70375	7.8270	10.7730	5.00	15.00
Total	60	9.5167	2.89647	.37393	8.7684	10.2649	5.00	15.00

As indicated in Table 7., three groups vary in standard deviation and standard error of mean. The mean of the three groups is different on the delayed post-test receptive vocabulary knowledge. To test whether this difference in retention of receptive vocabulary knowledge of three groups is significant or not the researcher conducted a One-way ANOVA on the delayed post-test of the receptive vocabulary knowledge. The results of the one-way ANOVA on the delayed post-test receptive vocabulary knowledge scores are shown in Table 8.

**Table 8.** One-way ANOVA: comparison between scores on post-test productive vocabulary knowledge scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	113.633	2	56.817	8.492	.001
Within Groups	381.350	57	6.690		
Total	494.983	59			

As the results of Table 8 show, an F ratio of  $F(2, 59) = 8.492$ ,  $p = .001 < .05$  indicated a statistically significant difference among the three groups; therefore, the third null hypothesis of the study was rejected. To find the location of the difference, a post-hoc LSD was carried out. The results are presented in Table 9.

The mean difference reported in Table 9. indicates a statistically significant difference between the online Quizlet flashcards (Experimental group 1) and control groups. There is also a significant difference between online Quizlet flashcards and student-created Quizlet flashcards (Experimental group 2) groups as presented in Table 9., the mean difference of  $M = 2.00$ ,  $p = .018 < .05$ . The most difference lies between the experimental group 1 and the control group with a mean difference of  $M = 3.35$ ,  $p = .000 < .05$ . Based on the mean differences, the three groups can be ordered in this way with regard to receptive knowledge scores on the delayed post-test: (1) online Quizlet flashcards (2) student-created Quizlet flashcards (3) control. Moreover, there is not a significant difference between the student-created Quizlet flashcards (Experimental group 2) and control group' scores on delayed pot-test receptive vocabulary knowledge as presented in Table 9. the mean difference of  $M = 1.35$ ,  $p = .104 > .05$ .

**Table 9.** LSD post-hoc test: comparison between delayed post-test receptive vocabulary knowledge scores

(I) group	(J) group	Mean Difference (I-J)	Std. Error	Sig.
Control	Experimental 1	-3.35000*	.81795	.000
	Experimental 2	-1.35000	.81795	.104
Experimental 1	Control	3.35000*	.81795	.000
	Experimental 2	2.00000*	.81795	.018
Experimental 2	Control	1.35000	.81795	.104
	Experimental 1	-2.00000*	.81795	.018

#### 4.1.4. Research hypothesis 4

H04. There is no significant difference between traditional method, Online Quizlet flashcards, and Student-Created online Quizlet flashcards on retention of productive vocabulary knowledge.

To evaluate the effectiveness of the treatment on retention of productive vocabulary knowledge and to test the fourth hypothesis, One-way ANOVA was run on participants' delayed post-test scores on productive vocabulary knowledge. The descriptive statistics of the post-test scores are shown in Table 10.

**Table 10.** Descriptive statistics for the delayed post-test productive vocabulary knowledge scores

Group	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
Control	20	4.9500	2.21181	.49458	3.9148	5.9852	2.00	12.00
Experimental 1	20	4.9000	3.19374	.71414	3.4053	6.3947	1.00	13.00
Experimental 2	20	5.3000	2.40832	.53852	4.1729	6.4271	2.00	12.00
Total	60	5.0500	2.60003	.33566	4.3783	5.7217	1.00	13.00

As indicated in Table 10., three groups vary in standard deviation and standard error of mean. The mean of the three groups is different on the delayed post-test productive vocabulary knowledge. To test whether this difference in retention of productive vocabulary knowledge of three groups is significant or not the researcher conducted a One-way ANOVA on the delayed post-test of the receptive vocabulary knowledge. The results of the one-way ANOVA on the delayed post-test productive vocabulary knowledge scores are shown in Table 11.

**Table 11.** One-way ANOVA: comparison between scores on delayed post-test productive vocabulary knowledge scores

	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	1.900	2	.950	.136	.873
Within Groups	396.950	57	6.964		
Total	398.850	59			

As Table 11. shows, an F ratio of  $F(2, 59) = .136, p = .873 > .05$  indicated no statistically significant difference among the three groups; therefore, the fourth null hypothesis of the study was accepted.

## 4.2. Findings

In this section, findings of the study are presented with reference to the research questions followed by their discussions.

### 4.2.1. Research question 1

Q1. Is there any significant difference between traditional method, online Quizlet flashcards, and student-created Quizlet flashcards in learning of receptive vocabulary knowledge?

The analysis of data collection for the study revealed that in terms of learning receptive vocabulary knowledge, the three groups of the study were significantly different. The difference in mean scores is proof to claim that the students who received the online Quizlet flashcards outperformed other groups. Based on these data and the analysis of SPSS, the first null hypothesis is rejected. In other words, there is a significant difference between online Quizlet flashcards and student-created online flashcards and control groups for Iranian EFL learners in learning receptive vocabulary.

The difference between the findings of the current study and previous studies was the difference in the scores of students in online Quizlet flashcards and student-created Quizlet flashcards groups. The students of online Quizlet flashcards significantly outperformed student-created Quizlet flashcards group. In a study conducted by Mojarradi (2014), there was not any significant difference between prepared flashcards and student-created flashcards in terms of vocabulary learning. In a study conducted by Hung (2015), there was not any significant difference between student-created flashcards and prepared flashcards in terms of vocabulary learning.

The first explanation for the difference in the results of the experimental groups and control group is the repetition. Vidal (2011) found that repetition is a major factor affecting learning words compared to the other factors like type of elaboration, type of words, and word form. Brown et al. (2008) found that words that repeated more often had a greater chance of being learned. The data revealed that many repetition of the words (20 to 50 times) will be needed for substantial leaning of words. The students of the control group just were taught the words in the classroom and they didn't have any chance for practicing and repetition. Quizlet website provides enough chance for students to repeat each word more than 5 times. Moreover, it enables students to review the words of a course again. One of the main reasons for the superiority of the two experimental groups over the control group is more repetition during the period. The students in the control group received the

words of each session in 30 minutes, while the students in the experimental groups learned the words gradually over a week with a lot of practice and repetition.

#### 4.2.2. Research question 2

RQ2: Is there any significant difference between traditional method, online Quizlet flashcards, and student-created Quizlet flashcards in learning of productive vocabulary knowledge?

The analysis of data collection for the study revealed that in terms of learning productive vocabulary knowledge, the students of the control group significantly outperformed the students in online Quizlet flashcards.

These findings and results contrast with some of the previous findings in regard to employing online flashcards to learn productive vocabulary knowledge. In a study conducted by Taghizadeh and Porkar (2018) use of online flashcards had significant effect on improving productive vocabulary knowledge of students. The reason of the contrasting might be related to the COVID-19 widespread in Iran. This caused a situation that the teacher was not able to check the experimental groups who were supposed to work on the flashcards. Some students might not devote enough time to practice the flashcards. The students of the control group participated in face-to-face class.

One of the important factors in learning a word is word consciousness. Word consciousness involves the awareness of word parts, word order, and word choice in different uses of the language, how words extend their meanings, underlying core meanings, and how words are learned. The learners should be encouraged to expand their knowledge about members of a lexical set. Using flashcards as an example of explicit learning cannot give enough information about productive usage of the words (Nation, 2001). Scott and Nagy (1997) argue for developing word consciousness with L1 learners and Graves (2006) also strongly confirms their opinion in this regard.

As mentioned above, one of the most important information that a learner should be taught about the words is how to use them in different contexts and sentences. Students in control group received some examples for each word and also informed about parts of speech related to each word. On the other hand, students in both experimental group due to the impossibility of adding examples in flashcards could not receive some examples related to each word. This factor can be a good reason for the lack of superiority of experimental groups in the productive parts of the tests. The findings of this study is in contrast with the study conducted by (Webb, 2007) that showed there is no significant difference between learning a decontextualized word and its translation and learning a word and its translation with a sentence context. Flashcards are not able to provide detailed information about different forms and usages of a words. Flashcards limit the chance of practicing different forms of the words. Vocabulary exercises should focus on internalizing the words not just practicing surface form-meaning level. Students need to deal with collocations and multiple-word units not just single words (Faraj, 2015).

#### 4.2.3. Research question 3

RQ3: Is there any significant difference between traditional method, online Quizlet flashcards, and student-created Quizlet flashcards on retention of receptive vocabulary knowledge?

The analysis of the data showed students of online Quizlet flashcards group significantly outperformed students of the student-created Quizlet flashcards and control groups. Based on this claim the third null hypothesis were rejected. This means the rate of forgetfulness was higher in control and student-created Quizlet flashcards groups.

Unfortunately, there were not any previous studies related to the effect of online flashcards on retention of the receptive vocabulary knowledge, so the researcher was not able to compare the results and findings of the current study with the studies conducted before.

Students in both experimental groups had many opportunities to review and repeat vocabulary. The strength of the connection between meaning and form will determine how students retrieve the meaning of the word when they see it or hear it or how to use it when they want to express a meaning. It is also important for students to have enough repetition in addition to the knowledge of form and meaning (Nation, 2001). There are some evidence that the knowledge of form-meaning from using flashcards can be remained for a long time (Bahrick et al., 1987; Beaton et al., 1995).

As mentioned in previous parts using flashcards can increase self-testing effect. Many studies proved that you will recall 50% more of learned information by testing the items you learned rather than spending the same amount of time for studying (Zung et al., 2022).

#### 4.2.4. Research question 4

RQ4: Is there any significant difference between traditional method, online Quizlet flashcards, and student-created flashcards in retention of productive vocabulary knowledge?

The results and findings of the study revealed that there was no significant difference between three groups of the study.

The results of this study are in contrast with some of the previous study. For instance, Alghamdi and Elyas (2020) claimed that the use of online flashcards had significant effect on vocabulary retention of students in Saudi Arabia. In another study conducted by Nemati (2009), the results of the study showed that use of flashcards significantly improve retention of productive vocabulary knowledge.

Testing and generating effects are two important factors in recalling words. Testing effect is more concerned with the initial test of words in the process of recalling words. Initial tests mostly include multiple choice and matching questions which are closely related to receptive knowledge of the words. Testing effect works in more limited situations. Generating effect is more concerned with free final testing of the words. As it can be understood the knowledge you need to recall in generating effect is more concerned with productive vocabulary knowledge (Karpicke & Zaromb, 2010). According to the lexical activation hypothesis, learner must search his or her semantic memory during the process of generation (Payne et al., 1986).

## 5. Conclusion

This study aimed at investigating the effect of online Quizlet flashcards and student-created online flashcards on learning and retention of receptive and productive vocabulary knowledge by Iranian EFL learners. The findings proved that both experimental groups outperformed the control group on the post-test and delayed post-test. Based on the reviewed literature and this small-scale experimental research, it can be concluded that teaching vocabulary needs to be updated based on the needs of the learners and tools that instructors have at their disposal. Too much emphasis on traditional methods can reduce the motivation and performance of learners. The new generation are digital native and the teacher should adapt their methods of teaching with the interests of their learners. Although learning through online flashcards has not been introduced recently, its use among Iranian teachers and language learners needs more attention.

## 6. References

- Afshari, M., Bakar, K. A., Luan, W. S., Samah, B. A., & Fooi, F. S. (2009). Factors affecting teachers' use of information and communication technology. *International journal of instruction*, 2(1). <https://doi.org/10.18848/1447-9494/cgp/v16i03/46189>
- Alghamdi, A. A., & Elyas, T. (2020). The Effect of Electronic Flashcards on EFL Students' Vocabulary Learning: The Case of Saudi Arabia. *Randwick International of Education and Linguistics Science Journal*, 1(2), 114-125. <https://doi.org/10.47175/rielsj.v1i2.79>
- Azabdaftari, B., & Mozaheb, M. A. (2012). Comparing vocabulary learning of EFL learners by using two different strategies: Mobile learning vs. flashcards. *The Eurocall Review*, 20(2), 47-59. <https://doi.org/10.4995/eurocall.2012.11377>
- Bahrack, H. P., & Phelps, E. (1987). Retention of Spanish vocabulary over 8 years. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 13(2), 344. <https://doi.org/10.1037//0278-7393.13.2.344>
- Basoglu, E. B., & Akdemir, O. (2010). A comparison of undergraduate students' English vocabulary learning: Using mobile phones and flash cards. *Turkish Online Journal of Educational Technology-TOJET*, 9(3), 1-7. <https://doi.org/10.24114/tj.v2i1.903>
- Beaton, A., Gruneberg, M., & Ellis, N. (1995). Retention of foreign vocabulary learned using the keyword method: A ten-year follow-up. *Second Language Research*, 11(2), 112-120. <https://doi.org/10.1177/026765839501100203>
- Brown, R., Waring, R., & Donkaewbua, S. (2008). Incidental vocabulary acquisition from reading, reading-while-listening, and listening to stories. *Journal of Education and Practice*. <https://doi.org/10.1080/17501229.2016.1203328>
- Faraj, A. K. A. (2015). Effective Strategies for Turning Receptive Vocabulary into Productive Vocabulary in EFL Context. *Journal of Education and Practice*, 6(27), 10-19.
- Farrokhi, F., Zohrabi, M., & Bolandnazar, A. (2021). The role of flashcard and wordlist strategies with concrete vs. abstract words on Iranian EFL learners' vocabulary development. *Iranian Journal of English for Academic Purposes*, 10(1), 88-107. <https://dorl.net/dor/20.1001.1.24763187.2021.10.1.6.0>
- Finocchiaro, M. (1986). English As a Second or Foreign Language. In: New York.

- Folse, K. S. (2006). The effect of type of written exercise on L2 vocabulary retention. *TESOL quarterly*, 40(2), 273-293. <https://doi.org/10.2307/40264523>
- Folse, K. S., & Briggs, S. J. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. University of Michigan Press Ann Arbor.
- Graves, M. F. (2006). The vocabulary book: Learning and instruction. Newark, DE: International Reading Association, 69-70.
- Hummel, K. M. (2010). Translation and short-term L2 vocabulary retention: Hindrance or help?. *Language teaching research*, 14(1), 61-74. <https://doi.org/10.1177/1362168809346497>
- Hung, H. T. (2015). Intentional vocabulary learning using digital flashcards. *English Language Teaching*, 8(10), 107-112. <https://doi.org/10.5539/elt.v8n10p107>
- Karpicke, J. D., & Zaromb, F. M. (2010). Retrieval mode distinguishes the testing effect from the generation effect. *Journal of Memory and Language*, 62(3), 227-239. <https://doi.org/10.1016/j.jml.2009.11.010>
- Khodashenas, M. R., Farahani, S. K., & Alishahi, Z. (2014). Flash cards versus animated cartoons: A comparative study in vocabulary teaching and learning. *International Journal of Language Learning and Applied Linguistics World*, 5(4), 290-298.
- Komachali, M. E., & Khodareza, M. (2012). The Effect of Using Vocabulary Flash Card on Iranian Pre-University Students' Vocabulary Knowledge. *International Education Studies*, 5(3), 134-147. <https://doi.org/10.5539/ies.v5n3p134>
- Köse, T., & Mede, E. (2016). Perceptions of efl learners about using an online tool for vocabulary learning in EFL classrooms: a pilot project in turkey. *Procedia-Social and Behavioral Sciences*, 232, 362-372. <https://doi.org/10.1016/j.sbspro.2016.10.051>
- Laufer, B. (1992). Reading in a foreign language: How does L2 lexical knowledge interact with the reader's general academic ability'. *Journal of Research in Reading*, 15(2), 95-103. <https://doi.org/10.1111/j.1467-9817.1992.tb00025.x>
- Laufer, B. (1998). The development of passive and active vocabulary in a second language: Same or different?. *Applied linguistics*, 19(2), 255-271. <https://doi.org/10.1093/applin/19.2.255>
- Lee, L. (2004). Learners' perspectives on networked collaborative interaction with native speakers of Spanish in the US. *Language Learning & Technology*, 8(1), 83-100.
- Mohammadnejad, S., Nikdel, H., & Oroujlou, N. (2012). Reactivating EFL Learners' Word Knowledge by Means of Two Techniques: Flashcards versus Wordlists. *International Journal of Linguistics*, 4(4), 393. <https://doi.org/10.5296/ijl.v4i4.2913>
- Mojarradi, S. (2014). The effect of using flashcards on ESL (English as a Second Language) students' ability to learn vocabulary. *International Journal of Scientific World*, 2(2), 56-61. <https://doi.org/10.14419/ijsw.v2i2.3291>
- Nation, I. S. (2001). *Learning vocabulary in another language*. Ernst Klett Sprachen.
- Nemati, A. (2009). Memory vocabulary learning strategies and long-term retention. *International journal of vocational technical education*, 1(2), 014-024. <https://doi.org/10.4314/marang.v20i1.56821>
- Nicholson, T. (1998). Teaching Reading The flashcard strikes back. *Reading Teacher*, 52, 188-192.
- Osman, S. N., & Al-Homoud, F. A. (2015). Flash Cards vs. Word Lists for Vocabulary Building among Majmmah High School Students. *Journal of Human and Administrative Sciences*, 201(3125), 1. <https://doi.org/10.12816/0024010>
- Payne, D. G., Neely, J. H., & Burns, D. J. (1986). The generation effect: Further tests of the lexical activation hypothesis. *Memory & Cognition*, 14(3), 246-252. <https://doi.org/10.3758/bf03197700>
- Richards, C.J. & Schmidt, R. (2002). *Dictionary of language teaching and applied linguistics*. New York: Pearson Education.





- Saslow, J., & Ascher, A. (2011). *Top Notch 1 Student Book and Workbook Pack* (3rd ed.). Pearson Education ESL.
- Scott, J. A., & Nagy, W. E. (1997). Understanding the definitions of unfamiliar verbs. *Reading Research Quarterly*, 32(2), 184-200.
- Schmitt, N. (2010). Key issues in teaching and learning vocabulary. In *Insights into non-native vocabulary teaching and learning* (pp. 28-40). Multilingual Matters. <https://doi.org/10.21832/9781847692900-004>
- Sitompul, E. Y. (2013). Teaching vocabulary using flashcards and word list. *Journal of English and Education*, 1(1), 52-58.
- Taghizadeh, M., & Porkar, P. (2018). Tablet, flashcard and SMS and their effects on EFL learners' attitudes and vocabulary knowledge. *way*, 6(01).
- Tan, A., & Nicholson, T. (1997). Flashcards revisited: Training poor readers to read words faster improves their comprehension of text. *Journal of Educational Psychology*, 89(2), 276. <https://doi.org/10.1037//0022-0663.89.2.276>
- Teng, F. (2018). Incidental vocabulary acquisition from reading-only and reading-while-listening: A multi-dimensional approach. *Innovation in Language Learning and Teaching*, 12(3), 274-288. <https://doi.org/10.1080/17501229.2016.1203328>
- Test, Q. P. (2001). Oxford University Press and University of Cambridge Local Examinations Syndicate. *Photocopiable at UCLES*.
- Tinkham, T. (1997). The effects of semantic and thematic clustering on the learning of second language vocabulary. *Second language research*, 13(2), 138-163. <https://doi.org/10.1191/026765897672376469>
- Verhallen, M., & Schoonen, R. (1998). Lexical knowledge in L1 and L2 of third and fifth graders. *Applied Linguistics*, 19(4), 452-470.
- Vidal, K. (2011). A comparison of the effects of reading and listening on incidental vocabulary acquisition. *Language Learning*, 61(1), 219-258. <https://doi.org/10.1111/j.1467-9922.2010.00593.x>
- Webb, S. (2007). Learning word pairs and glossed sentences: The effects of a single context on vocabulary knowledge. *Language teaching research*, 11(1), 63-81. <https://doi.org/10.1177/1362168806072463>
- Zung, I., Imundo, M. N., & Pan, S. C. (2022). How do college students use digital flashcards during self-regulated learning?. *Memory*, 30(8), 923-941. <https://doi.org/10.1080/09658211.2022.2058553>



<https://doi.org/10.22077/ali.2023.7268.1029>

# 112 Years of Medical English: A Scientometric Analysis

Narges Hekmati<sup>1\*</sup>   
Ali Behdani<sup>2</sup> 

<sup>1</sup>Assistant Professor in Applied Linguistics, Department of General Courses, School of Medicine, Birjand University of Medical Sciences, Birjand, Iran

<sup>2</sup>M.A. Student in Applied Linguistics. Department of English Language, Faculty of Literature and Humanities, University of Birjand, Birjand, Iran

## ABSTRACT

This study addresses the science mapping and visualization of the 112 years of academic literature on medical English. The data for the present study was retrieved from the Web of Science database, which contains all English-only articles from 1912 to 2024. The scientometric techniques and analysis were done using VOSviewer. These methods of analysis include publication and citation patterns, co-authorship and co-occurrence networks, and bibliometric coupling of items. The findings indicated that the total number of articles published in Medical English from 1912 to 2024 was 10396. The publications surged from 1988 to 1992 and had a steady rise until its peak in 2023. Citations, however, differ from publication trends and have fluctuated during this time frame. Furthermore, authors, institutions, and the country's collaborative networks were examined to have a snapshot of the relationships across disciplines. The findings revealed a strong correlation between co-authorship and bibliometric coupling of countries, which shows that the USA, the UK, Australia, and Canada collaborate the most within the literature. Universities of San Francisco, Toronto, and Washington are considered among the leading research institutions in terms of output, while universities such as San Francisco, Harvard University, and Toronto rank among the top in terms of citations. Moving to author-level metrics, the output and citation patterns indicate that De Lusignan and Schillinger are the most prolific authors. At the same time, Shimada, Yoshida, and Grumbach have the most citations. Moreover, keyword co-occurrences showed that keywords such as "care," "impact," "healthcare," "health," and "quality" tended to appear most frequently in the literature.

**KEYWORDS:** Medical English; Scientometric analysis; English for specific purposes (ESP); Bibliometric review; Science mapping

## ARTICLE HISTORY

Received: 14 May 2023  
Revised: 09 July 2023  
Accepted: 10 August 2023  
Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [nargess.hekmati@gmail.com](mailto:nargess.hekmati@gmail.com)

## 1. Introduction

According to Hutchinson and Waters (1987), English for Specific Purposes (ESP) emerged in the late 1960s. They state three main developments underlying the emergence of ESP: "the demands of a Brave New World," "a revolution in linguistics," and "focus on the learner." Specifically, after World War II, there was significant worldwide expansion in science, technology, and economics led by the United States, which created a need for many people to learn English not for leisure but to meet practical needs. The second factor, as they put it, was a shift in linguistics from studying language usage and grammar to studying actual language use in communication, which raised awareness of differences in the English needed for different fields. Finally,

changes in educational psychology led to a shift from teacher-centered to learner-centered approaches focusing more on students' needs and interests (Hutchinson & Waters, 1987).

The emergence of ESP was based on recognizing that language learners have varying needs, and identifying them carefully should underpin the design of any syllabi or materials. Paltridge and Starfield (2013) state that analyzing students' language needs lies at the core of ESP. Therefore, many researchers have underlined the crucial role of needs analysis in developing ESP courses, including Bachman and Palmer (1996), Dudley-Evans and St John (1998), Hyland (2006), Long (2005), and Richards (2001).

Initially, ESP focused mainly on technology and business (Benesch, 2001). However, its scope has expanded over time to cover other areas like English for occupational, academic, vocational, business, legal, sociocultural, and medical purposes (Belcher, 2009). Discussing English for Medical Purposes (EMP), Salager-Meyer (2014) traces its origins to 1983 when cross-pollination between ESP and health sciences was first evidenced at a conference which led to the launch of the first EMP publication, the "EMP Newsletter" journal.

While proficiency in English as the international language is vital across scientific and technical fields, it is especially imperative for students pursuing medical sciences. This is not only due to the predominance of English-language publications but also the rapid advances and frequent updates in medical knowledge. Medical students must continually access the latest sources to ensure they will become effective practitioners. Thus, understanding English is crucial for keeping up with the swift pace of change in the field and applying cutting-edge information to enhance patient care.

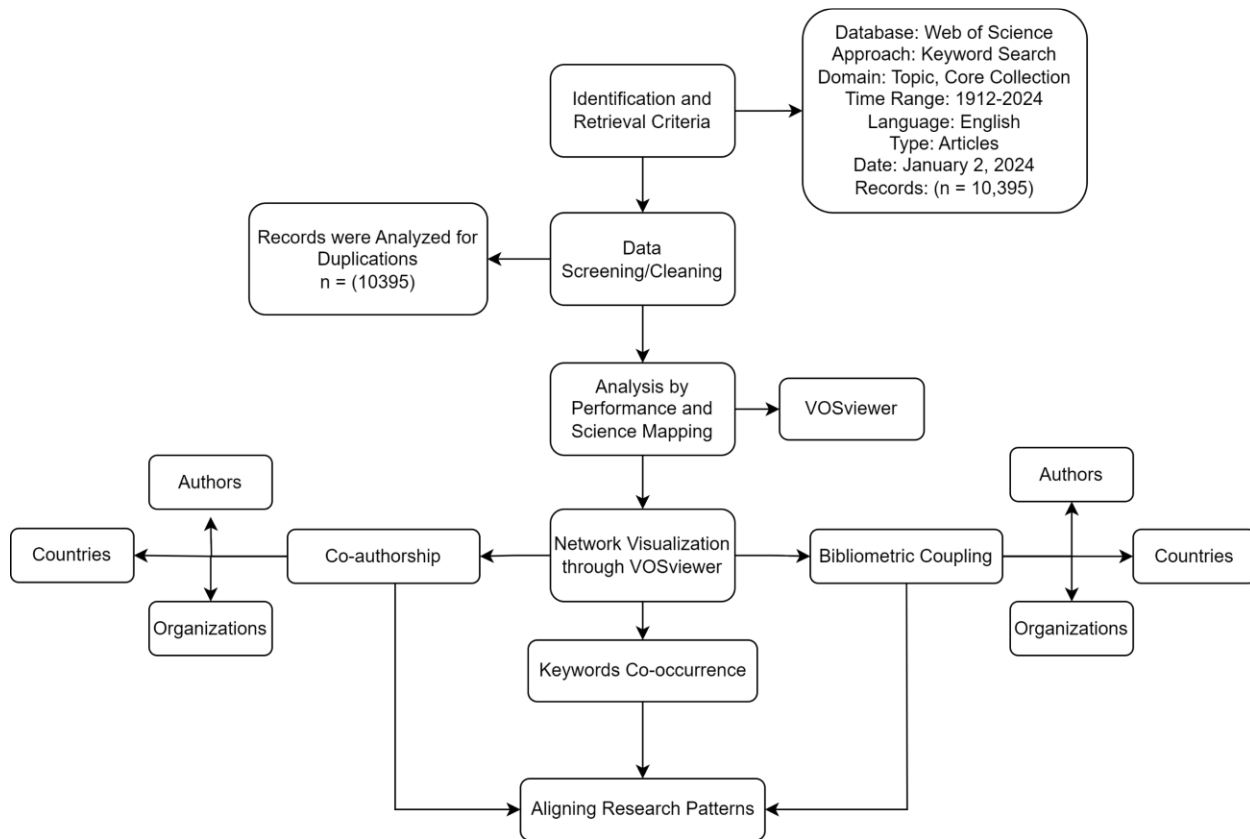
Given the importance of medical English, education scholars have endeavored over the past century to advance understanding of this research area. Conducting a bibliometric analysis of these works would provide valuable meta-level perspectives into publication trends, influence dynamics, and other dimensions within this substantial body of literature. Examining factors such as production volume, citation counts, co-authorship networks, institutional flows, geographic dispersion, topical evolution, and more offers crucial insights. Similarly, highlighting the role of particular nations, journals, and authors spotlights pivotal framing hubs. Additionally, tracing bilateral citations can uncover underappreciated contributions linking different language-teaching communities. Gaining such comprehension of the epistemic landscape around medical English assists researchers, experts, and reformers in positioning their efforts for optimal impact. The significance of this bibliometric study on medical English is compounded by the need for prior inquiries directly examining this specific domain.

Therefore, the present study aimed to thoroughly review and synthesize the vast scholarly works focused on medical English published over 112 years. This study aimed to provide enhanced clarity surrounding the evolutionary trajectory that has defined research developments within this domain across the twentieth and early twenty-first centuries by comprehensively examining this extensive body of literature. In particular, detailed analysis targeted elucidation of underlying patterns, trends, and shifts in the foci and methodological orientations of inquiry over time to reveal a coherent portrait of how the investigation into medical English has progressed. Mapping the complex publication topography further highlighted remaining gaps and unresolved issues, meriting additional research attention.

## 2. Methodology

This study utilized a keyword search in the "Topic" domain of the Web of Science (WoS) core collection. The term was "Medical English," the results were then filtered to English-only articles. This method returned 10,395 documents. The descriptive statistics tools available in WoS showed that within these documents, there are metadata for 43,102 authors, 8,787 affiliations, 3,206 sources, 872 publishers, 130 research areas, 156 countries, 207 WoS categories, and 4,444 funding agencies. The period these datasets were a part of covered an astounding interval of 112 years from 1912 until January 2, 2024. The bibliographic data was then analyzed using bibliometric tools. VOSviewer (Eck & Waltman, 2010) was utilized to conduct co-authorship, bibliometric coupling, and co-occurrence network visualizations. Tackling these visualizations can provide further insight into the collaboration patterns between authors, the relationship between keywords, and the structure of medical English as a research field.

Moreover, network visualizations of VOSviewer present a handful of information for us to analyze, such as clusters and prominent nodes that represent the main topics, collaborations, and impactful publications, overlay visualizations, and density maps, which are a part of analyzing the data based on frequency and time. These features are addressed in the result and discussion section. Figure 1 shows a thorough visualization of the methodology.



**Figure 1.** Data selection process and methodology review

## 2.1. Scientometric terminologies

Scientometric techniques differ from one study to another. Some tend to analyze data themselves, whereas others utilize applications. However, the benefit of applications surpasses those doing the analysis themselves. VOSviewer is one of those applications with a handful of measurement and analysis techniques delving deep into the dataset. Another benefit of using VOSviewer is the eye-catching visualization that comes with the analysis. The measurement techniques brought to bibliometric and scientometric analysts include co-authorship, co-occurrence, bibliometric coupling, references, citation patterns, and co-citation patterns across the dataset. Each technique can be used on a specific group, e.g., countries, authors, organizations, etc.

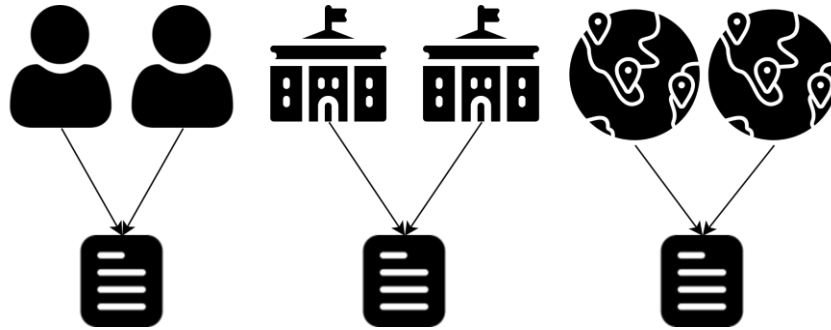
This study utilized three measurement techniques for all groups: co-authorship, co-occurrence, and bibliometric coupling. Along these approaches, there are other methods, which we will discuss later in the section. To comprehensively analyze the visualization data, we briefly explain the terms to avoid confusion.

Co-authorship refers to two or more groups (countries/authors/organizations) writing a document within the literature together (Ullah et al., 2022). For example, the co-authorship of the United States is an "N" number, which means the United States has had "N" documents co-authored in the literature.

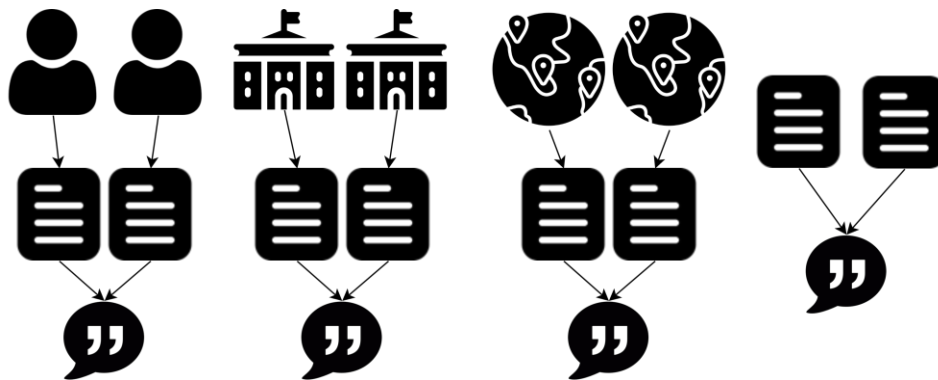
Co-occurrence is another measurement used for these types of studies. It focuses on the counts of co-occurring entities in a text corpus (Zhou et al., 2022). However, this technique is only used to analyze keywords present within the dataset. There are two keywords: Author Keyword and Keyword Plus. Author keywords, as the name shows, are keywords authors give to their paper, which is more specific but limits the analysts to study the trends and thematic evolution. On the other hand, keywords plus are frequent words and phrases present in the title of the references but do not appear in the article's title. This group has broader terminology and is better for bibliometric studies based on Zhang et al.'s (2015) findings.

According to Kessler (1963), Bibliometric coupling measures how similar documents are based on their shared references. This type of measurement can find the relationship between the documents, countries, and authors.

For a better understanding of these terms, the following flowcharts are provided which are designed using the images from flaticon.com (Figure 2 & 3).



**Figure 2.** Co-authorship of authors, organizations, and countries



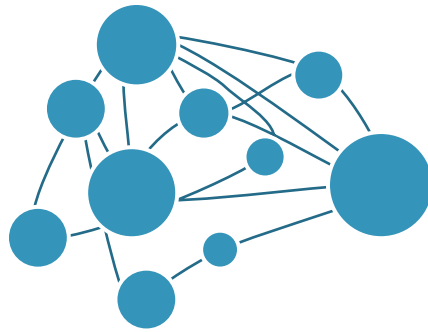
**Figure 3.** Bibliometric coupling of authors, organizations, countries, and documents

Beside the general terminology, we need to understand the specific terminologies for each map that VOSviewer creates. These terms include Nods, Links, and Clusters.

A nod is a circle in the network presenting an idea, a theme, an author, an organization, a country, a document, or a reference. If this nod gets bigger it means the frequency is more significant compared to other nods.

A link is the tile or the connection between the nods. It can be spotted by looking for interconnected curved lines between the nods. These links show the relationship of the nods with each other.

Clusters are groups of people, organizations, references, documents, and so on, which can be presented in different colors in the network. An example of the terms in one setting could be a dark blue cluster that has 10 authors which are presented by nods with different sizes indicating the frequency. These nods then are connected with links that show the relation between them. Figure 4 is presented for better comprehension of the terms.



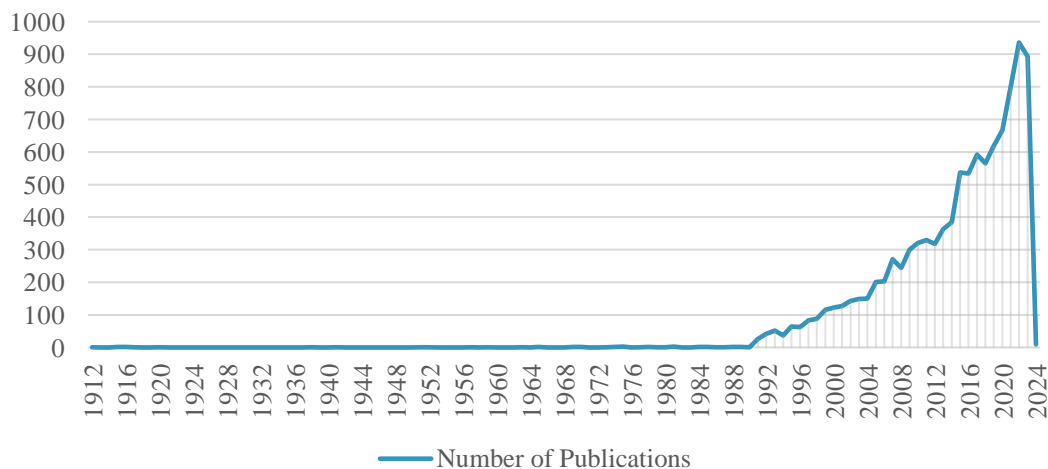
**Figure 4.** Nods, links, and clusters

### 3. Results and discussions

Having introduced the terminologies in the previous section, we now move to the result and discussion sections of the study. The findings are presented separately based on the groups (Authors, Organizations, Countries, and Keywords). Moreover, in each section, different methods were used to visualize and comprehend the dataset.

#### 3.1. Overview

In this section, we address the journal-level metrics, which include the number of publications/citations and the most cited documents. Figure 5 indicates the number of publications from the early decades. As shown in the figure, annual scientific publication for medical English was relatively infrequent, with only two or three papers published annually up to the late 1980s. In the 1990s, production had a slow upward trend, with about 50 papers published annually. The publication rate grew with a significant number of documents, well over 100 papers per year in most years. In the past decade, productivity has increased dramatically, from 300 papers per year in 2010 to over 800 by 2021. Consequently, medical English research attention and output have grown exponentially in recent years.



**Figure 5.** Annual scientific production

Scientific production mirrors the findings for citation patterns (Figure 6), which showed an upward trajectory from the beginning of the research in the early 20th century to 2007. Papers from 1912 to 1989 had an average citation of 0 to a little over 0.5. However, with the start of the year 1990, the average number of citations of articles began to rise, with fluctuations in between. This upward trend continued up until 2007, which was the peak average citation of articles in the medical English discipline. From 2008 forward, a downward trend started with plenty of fluctuations, finishing in 2023 with an average citation

of less than 0.5. Based on these findings, it is clear that the research conducted during this period laid essential foundations and had a more significant influence on scholarly discourse. However, since 2007, the average number of citations per paper has declined to around 2 per year. Furthermore, the recent increase in publications has yet to be translated into a higher impact on citations.

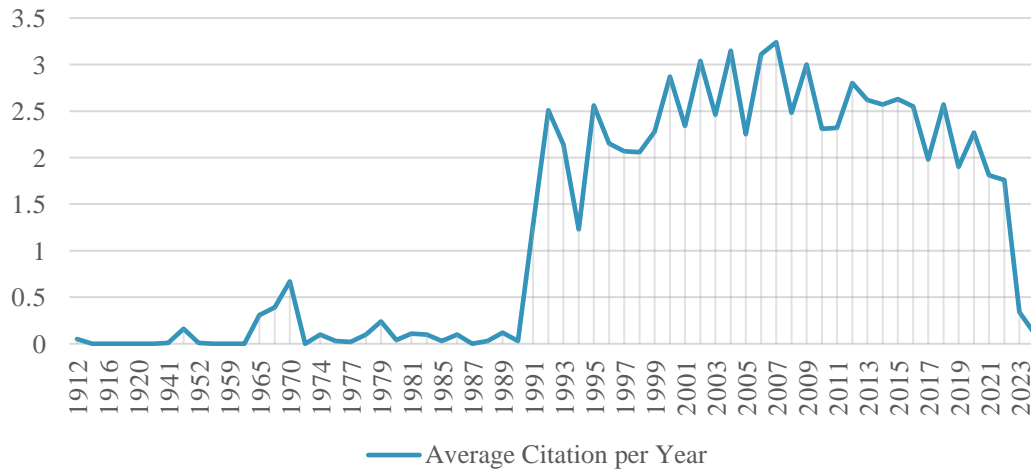


Figure 6. Average citation per year

Aside from the patterns of publication and citation for the discipline, addressing the most cited documents can provide further insight into the literature and help us delve deeper to find the core ideas related to the field. Notably, many of these highly impactful papers (Table 1) relate to health communication, medical education, heart diseases, and heart failures, as well as improving outcomes through enhanced provider-patient relationships. For example, the top-cited article by Merikangas et al. (2007) highlights the prevalence of bipolar spectrum disorder nationally. Additionally, Schillinger et al. (2002) and Jack et al. (2009) address research areas such as health literacy and discharge programs/rehospitalization. Other highly-cited papers are tackling heart failure, education in medicine, staffing patterns, clinical outcomes, shared decision-making, and neurodevelopmental outcomes.

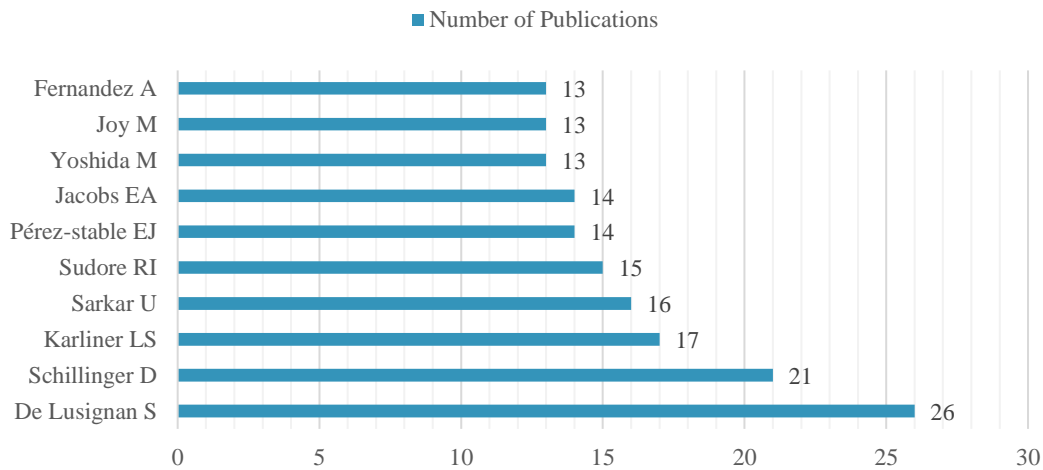
Table 1. Top ten most-cited articles

Paper	DOI	Total Citation
Merikangas Kr, 2007, Arch Gen Psychiat	10.1001/archpsyc.64.5.543	1669
Schillinger D, 2002, Jama-j Am Med Assoc	10.1001/jama.288.4.475	1181
Jack Bw, 2009, Ann Intern Med	10.7326/0003-4819-150-3-200902030-00007	1103
Antman Em, 1992, Jama-j Am Med Assoc	10.1001/jama.268.2.240	1064
Pronovost Pj, 2002, Jama-j Am Med Assoc	10.1001/jama.288.17.2151	1043
Frank Jr, 2010, Med Teach	10.3109/0142159X.2010.500898	1035
Makoul G, 2006, Patient Educ Couns	10.1016/j.pec.2005.06.010	1009
Marino Bs, 2012, Circulation	10.1161/CIR.0b013e318265ee8a	1006
Heidenreich Pa, 2022, Circulation	10.1161/CIR.0000000000001063	986
Goldstein Dj, 1992, Int J Obesity	NA	979

### 3.2. Authors

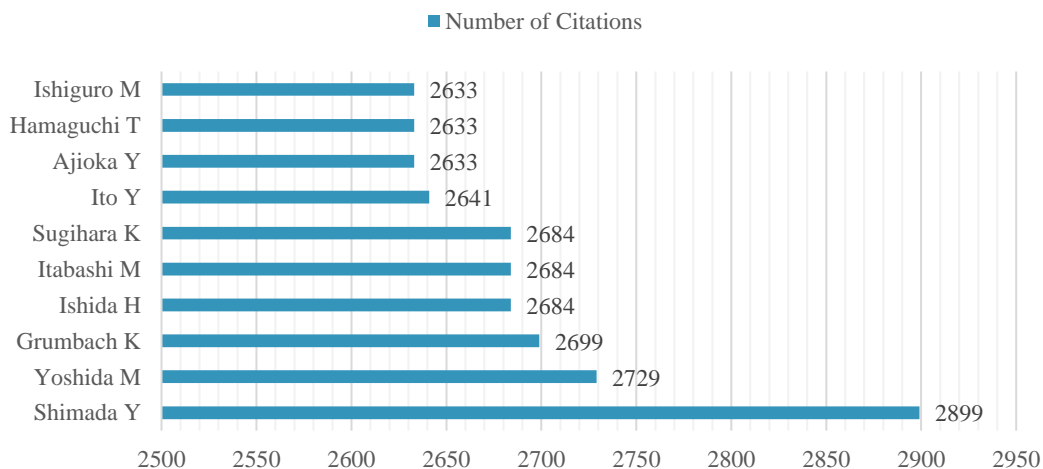
Author-level metrics include the most cited and published authors within the literature and co-authorship and bibliometric coupling of the writers who contributed to the field of medical English.

Identifying central authors and analyzing their co-authorship structures sheds light on patterns of ideation, team dynamics, and influence shaping the medical English field. Figure 7 shows De Lusignan S with 26 papers, followed by Schillinger D with 21 as top authors. The other top authors include Karliner LS, Sarkar U, Sudore RI, and Pérez-Stable EJ, with 14-17 publications each.



**Figure 7.** Top ten most-published authors

While prolific authors contribute extensively to the volume of medical English literature, citation data (Figure 8) reveals those producing the most impactful scholarship. A total of 2,899 citations are reported by Shimada Y, indicating the researcher's significant contribution to the field. Following Yoshida M and Grumbach K, both have accumulated substantial impact through their work. Several highly cited Japanese authors, including Ishida H, Itabashi M, Sugihara K, and Ito Y, each with citation counts of approximately 2,600, indicate a highly cited regional research community. Ajioka Y and Hamaguchi T. further demonstrate Japan's strength in influential scholarship. It is apparent from comparing citations and publication leaders that productive authors contribute to the foundation of literature. However, less prolific yet highly insightful authors may generate genuine field-shaping insights.

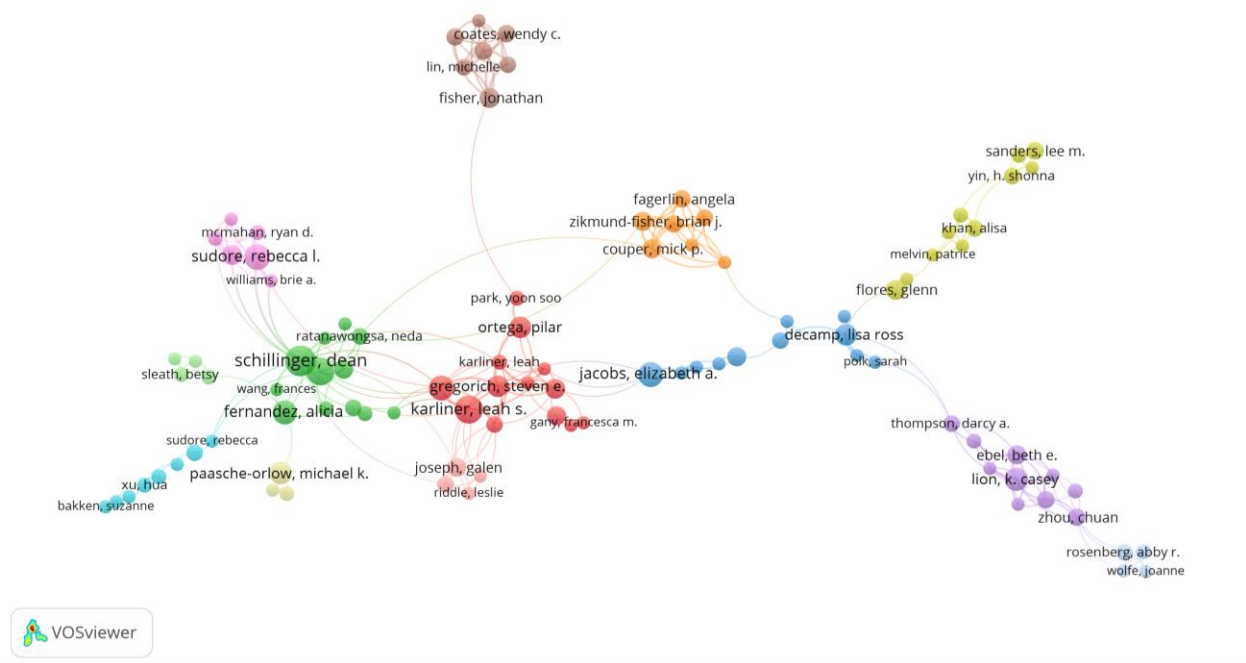


**Figure 8.** Top ten most-cited authors

Complementing the citation analysis, co-authorship networks among authors were also examined further to reveal collaborative patterns within the medical English literature. Figure 9 visualizes the co-authorship network of the top 100 most

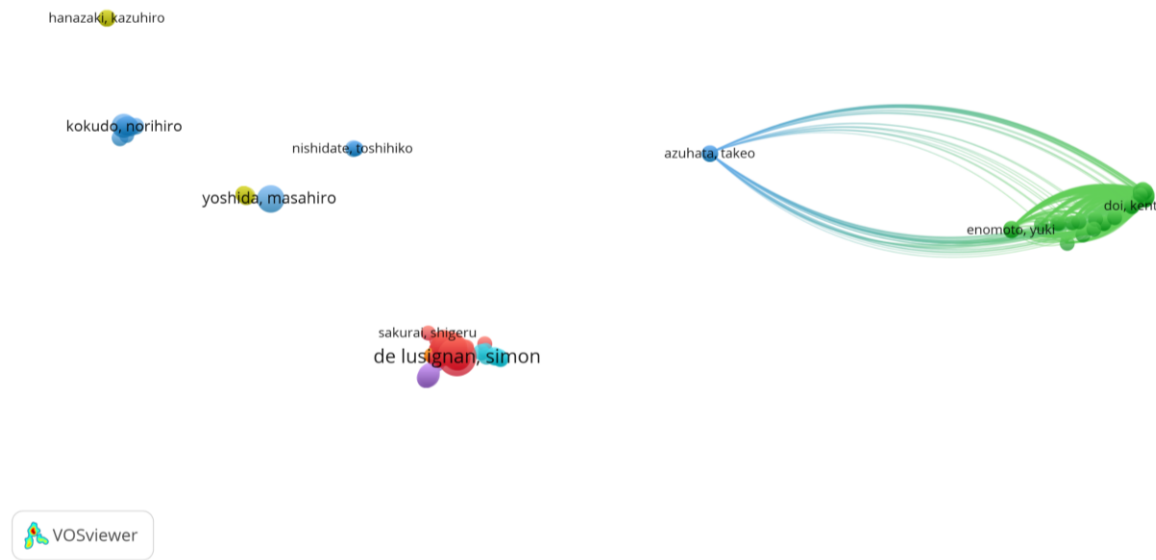


connected authors, clustered into 13 components linked by 243 ties. The most significant clusters in red, green, and dark blue contain 13, 12, and 11 members, respectively, representing the densest collaborative groups. Research communities centered around shared interests and frequent interaction between members can be identified when mapping these co-authorship structures. The degree of cooperation and isolation in knowledge production can be determined by examining the connectivity within and between clusters. Highly interconnected clusters indicate a deep collaboration and exchange of ideas within a particular subject area. Sparse connections between separate components indicate fragmentation of the field. As these co-authorship network patterns evolve, we can determine how team dynamics and research networks evolve in medical English scholarship.



**Figure 9.** Authors’ co-authorship network

Furthermore, as mentioned before, bibliographic coupling analysis identifies the connections between published documents based on shared references. An in-depth examination of coupling relationships is an excellent method of discovering the links that hold a research field together. The bibliographic coupling network visualized in Figure 10 contains 541 of the most connected papers clustered into seven components. The largest red cluster accounts for nearly three-quarters of all items, representing a dense core of related literature. The second and third places go to green and blue clusters containing 43 and 38 members, respectively, signifying smaller secondary groupings. The predominance of the central red cluster suggests a unified conceptual framework with many papers building on overlapping references.

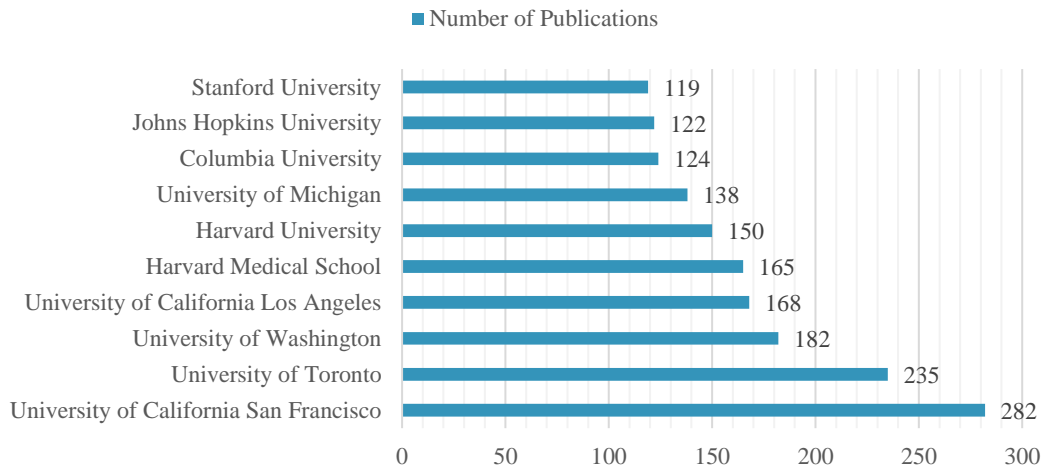


**Figure 10.** Authors' bibliometric coupling network

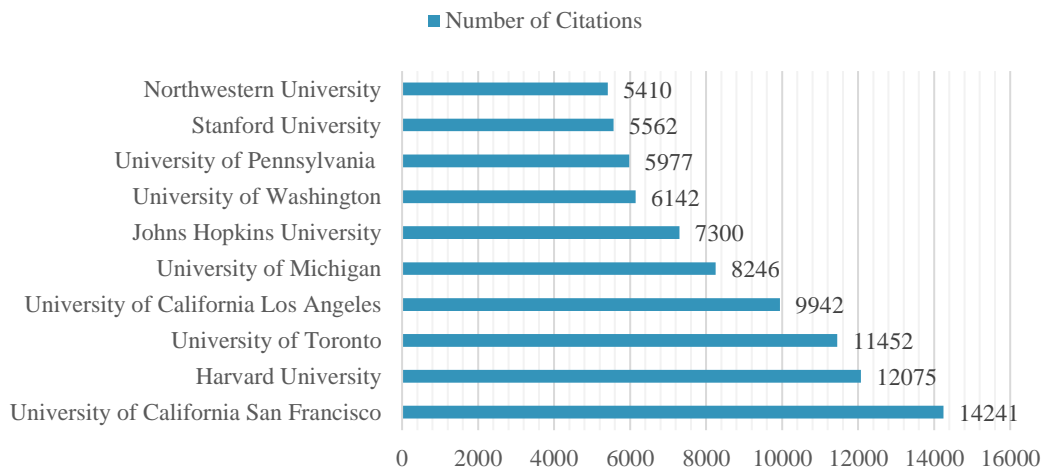
### 3.3. Organizations

In the realm of scientometric studies, organizational metrics remain a prominent aspect, giving an overview of the institutional research landscape. Figure 11 is provided to visualize a part of this landscape. The University of California San Francisco leads in medical English research output with 282 total publications, showing the institute's significant contribution. The University of Toronto ranks second with 235 papers, emerging as a critical research center in Canada. The University of Washington places third with 182 publications, representing rising productivity from this establishment. Other top producers include the University of California Los Angeles (168 papers), Harvard Medical School (165), Harvard University (150), and the University of Michigan (138), reflecting ongoing scholarship from these established research hubs. Moreover, Columbia University (124 papers) and Johns Hopkins University (122) are emerging as new research organizations. Rounding out the top 10 is Stanford University, with 119 publications.

Matching the data in output patterns, the citation pattern (Figure 12) of The University of California San Francisco again tops the list with 14,241 citations in the data analysis, indicating the highly influential nature of the university, its policy, and the researchers contributing to the literature through the institute. Harvard University holds the second place with 12075 citations, followed by one level demotion of The University of Toronto with 11452 citations in total. The University of California Los Angeles, the University of Michigan, and Johns Hopkins University also make the top citation list, with 9942, 8246, and 7300 citations, respectively. Finally, The University of Washington, The University of Pennsylvania, Sanford University, and Northwestern University are the top ten most-cited organizations with 6142, 5977, 5562, and 5410 each. Identifying organizations with a high or low influence can be achieved by comparing citation counts to publication numbers. A few examples include Northwestern University and the University of Pennsylvania, which are ranked among the top ten most-cited organizations but need to appear in the top ten most-published organizations, suggesting their contribution to have lower output at a higher level.

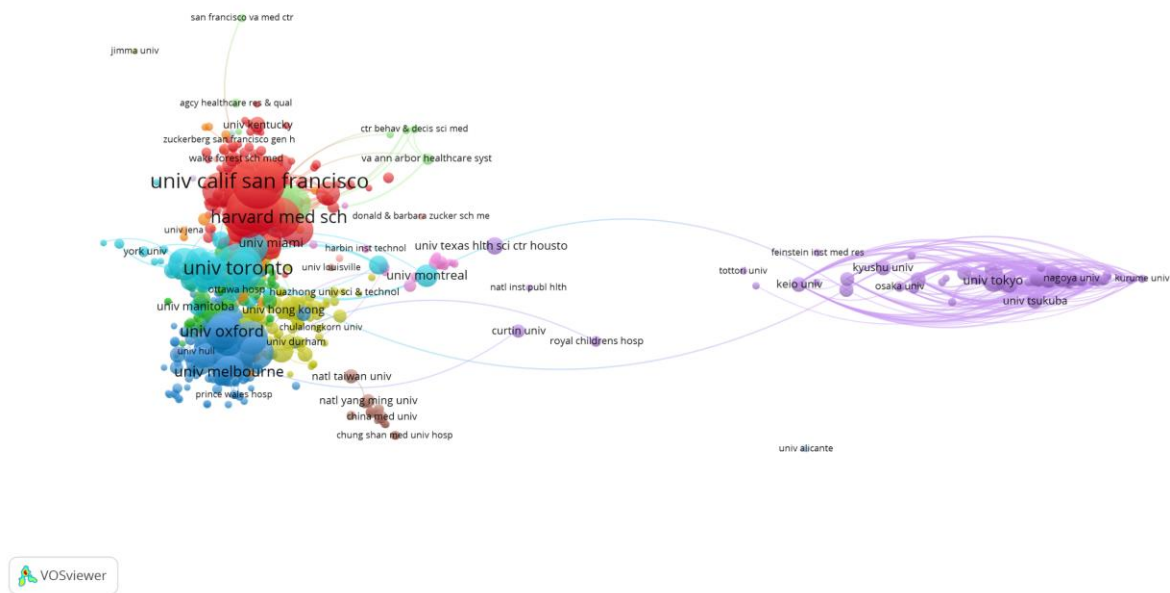


**Figure 11.** Top ten most-published organizations

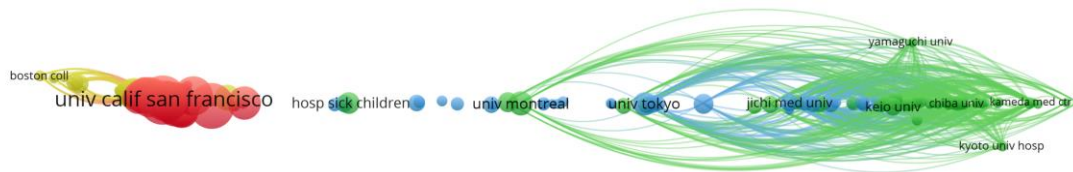


**Figure 12.** Top ten most-cited organizations

In addition to publication and citation patterns in the mentioned figures, Figure 13 visualizes the dense collaboration network between organizations researching medical English globally. This map contains 937 institutional nodes clustered into 15 components. The nodes present in these clusters are connected with 13,730 links. In addition, the largest cluster colored in red contains 233 items, followed by the green and dark blue clusters with 173 and 172 members, respectively. According to the figure, the mentioned components have a high degree of collaboration and knowledge exchange within the literature. The smaller clusters with lower member counts represent niche topics and research interests.



**Figure 13.** Organizations' co-authorship network



**Figure 14.** Organizations' bibliometric coupling network

Looking at Figure 14, we can see a dance coupling among the institutional networks. This visualization, with 946 institutional nodes clustered into four components connected by 184,134 links, shows significant coupling in the literature. The most dense cluster colored in red contains 854 items, followed by the second-order green and dark blue clusters in third place, having just 45 and 32 members, respectively. The significant size of the clusters shows the amount of coupling they have within the dataset. Larger clusters show the overlap of references within the documents, and smaller-sized clusters show differing references.

Analyzing the networks and visualizations provided through maps and charts can provide a deeper insight into the breadth and insularity of cooperation in medical English research worldwide. This can let researchers discover the changing dynamics and institutional interactions and tackle further investigation for emerging trends.

### 3.4. Countries

A highly debatable topic in scientometric domains is country metrics, which can provide convenient data for scholars worldwide. These data metrics can explain why one country contributes more to a discipline than others. Sometimes, this matter is self-explanatory and does not need further elaboration. However, this is not the case in some fields, and addressing these issues can lead to eye-opening results. Furthermore, these results can mirror the publication patterns (what to publish, when to publish, where to publish) of top countries in the lower rank countries regarding publications and collaborations.

Moving to publication output patterns for medical English (Figure 15), the United States leads the list with 4,312 papers. The UK is in second place with 1,476 publications, followed by Canada (817), Australia (682), and emerging output from China (539). Notably, the presence of English-speaking countries continues. However, the landscape shows a broader

global participation. For instance, European countries, including Germany (283 papers), Italy (251), Spain (242), and France (218), demonstrated expansion of contribution, along with Japan (275) representing Asia. Looking at the results, the geographic distribution has become more balanced, though North America and Western Europe still account for the majority of outputs.

The United States also leads in total citations (Figure 16) with 151,249, followed by the UK, which again holds second place with 37,535 citations. Matching the two figures for Figures 12 and 13, we can see that alongside a high output rate, citation rates also top the list, indicating a high volume of impactful work for these two countries. Canada follows with 29,625 citations; however, its citation number exceeds its publication output, indicating success in publishing impactful scholarship. Australia and Japan also perform well on both output and citation measures. Emerging countries like China (6,919 citations) have notably lower citation rates than output. Comparing citation rates reveals that non-English speaking countries like Switzerland, Italy, Germany, and the Netherlands have more influential research in terms of citation than output.

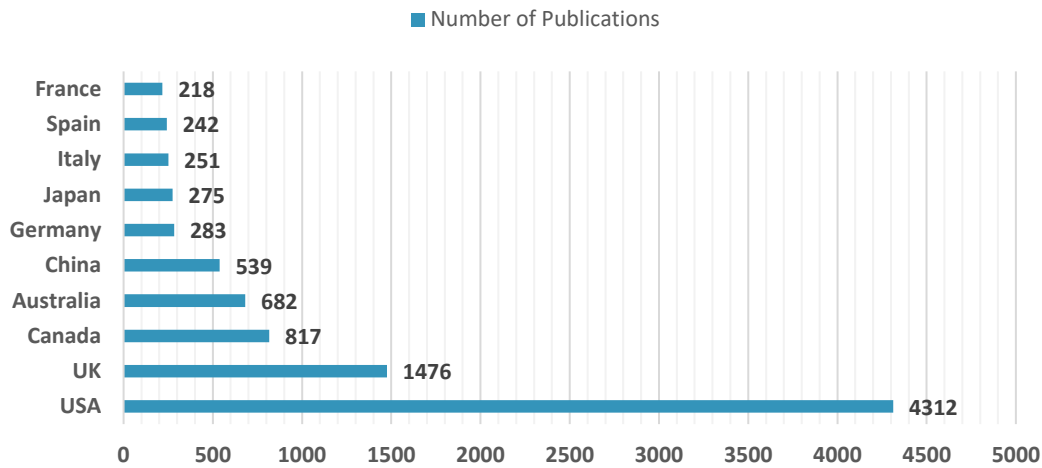


Figure 15. Top ten most-published countries

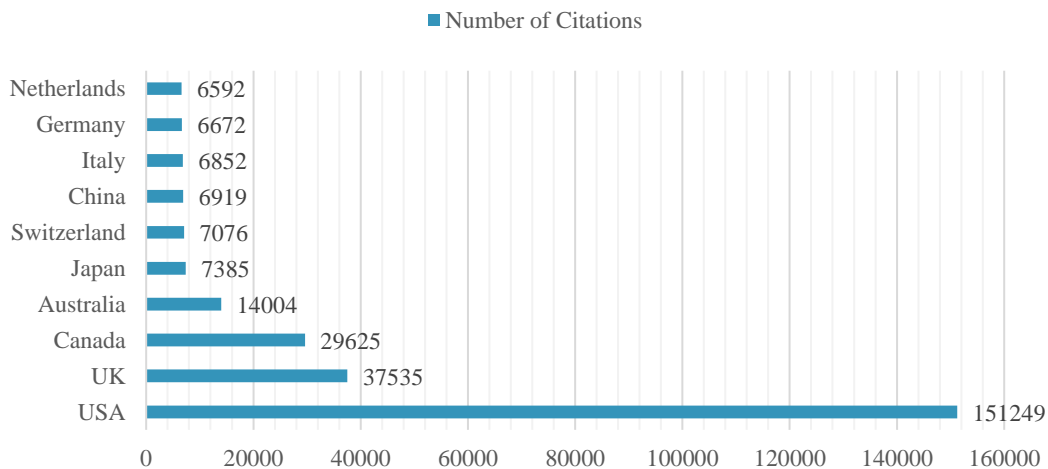
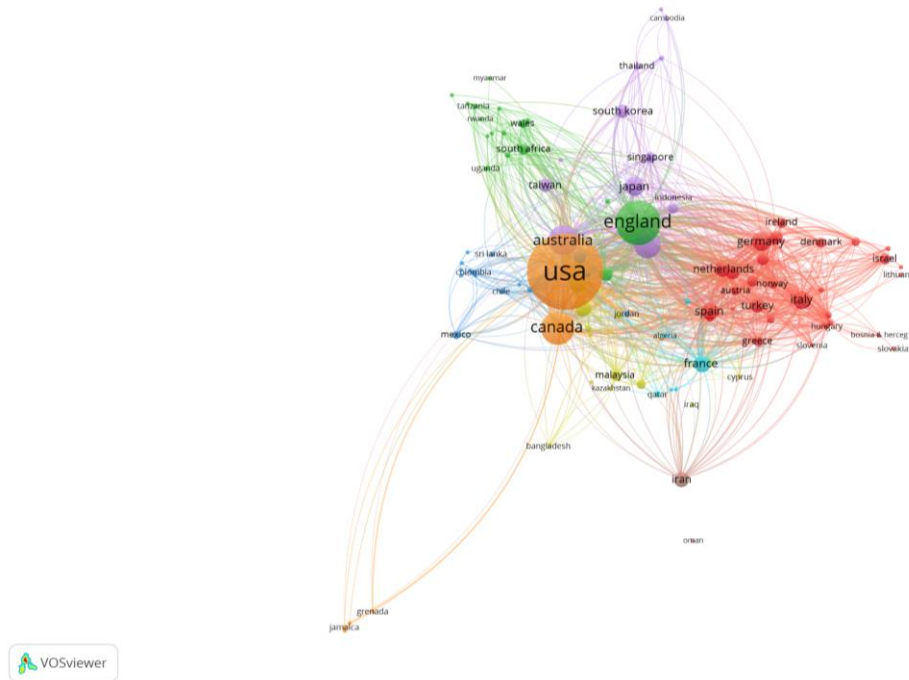


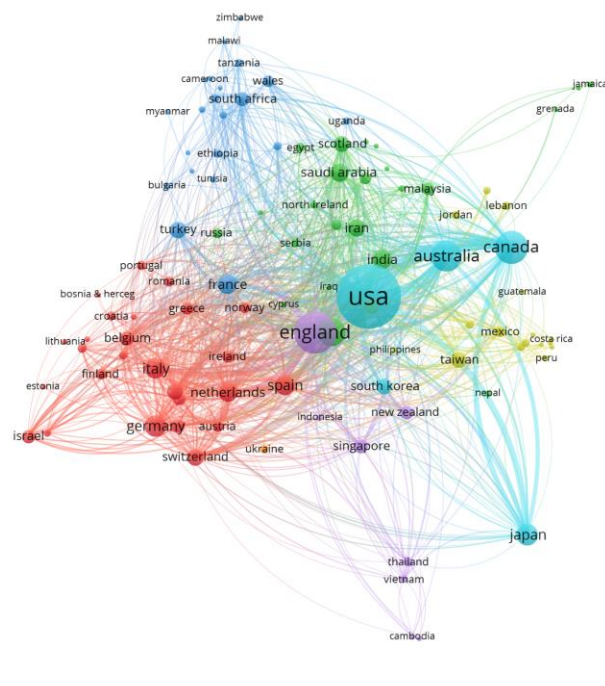
Figure 16. Top ten most-cited countries

Using VOSviewer, we generated the collaboration and coupling network of the countries. Figure 17 visualizes the dense relations between the countries in publishing scholarly output. The map comprises 102 country nodes clustered into eight components, with 1,766 links between them. The largest cluster colored in red contains 29 items, followed by the green and dark blue clusters, which comprise 17 and 15 members, respectively. The largest nodes present on the map are the USA, England, Canada, Australia, and France. These significant components represent regional collaborative groupings, with densely interconnected links reflecting close cooperation between nations.

Figure 18 shows the bibliometric coupling relationships between the countries. Similarly, this analysis revealed 107 country nodes clustered into seven components, connected by 3,849 coupling links. The most significant clusters colored in red, green, and dark blue contain 26, 24, and 21 members, respectively. As with the co-authorship network, the United States, England, Canada, and Australia are critical nodes across all clusters, pointing to their broad influence.



**Figure 17.** Countries' co-authorship network



**Figure 18.** Countries' bibliometric coupling network

### 3.5. Themes and keywords

Moving to themes analysis, the authors provided five different networks alongside the frequency of top-occurring themes in the dataset. These visualizations provide comprehensive insights into the thematic evaluation and hot/cold topics across time. To better decode the networks, the maps were created using different color palates to ensure maximum clarity for the readers; before going into the analysis of the networks, we first address the most-occurring keywords across the dataset. Scientific and analytical studies heavily rely on keywords and thematic analysis as part of the scientometric analysis. Visualizing these keywords can delve into the core, past/present, and future themes, which can then be compared and contrasted to eventually indicate the evolution of themes. Table 2 shows the top 20 most-occurring keywords in the entire article. The variables for this analysis were Authors' Keywords and Keywords Plus. Authors' Keywords are exactly what the names suggest: the keywords authors give to their work. These keywords can sometimes be limiting because they represent the core themes. Keywords Plus, however, are more general than authors' ones. These keywords occur in the titles of the documents' references in the dataset but not in the documents' titles.

This analysis revealed that the most frequent terms are broad keywords like "care," "health," "quality," "management," and "outcomes," followed by other top keywords including "education," "language," "communication," "disparities," and "United States." The prominence of terms like "women," "children," "risk," "cancer," and "diagnosis."

**Table 2.** Top 20 most frequent keywords

Keyword	Frequency	Keyword	Frequency
care	770	education	350
impact	529	language	343
health-care	501	women	338
health	501	risk	326
quality	481	disparities	315
management	467	united-states	309
outcomes	436	quality-of-life	294
communication	407	cancer	270
prevalence	406	validation	270
children	400	diagnosis	264

(Authors' Keywords & Keywords Plus)

Visualizing the keyword co-occurrence revealed a dense connection between the thematic and knowledge evolution of medical English literature. This network visualization (Figure 19) shows 2,614 terms clustered into 12 components, with 116,936 links between them. The three most significant clusters colored in red, green, and dark blue contain 643, 635, and 242 members, respectively. These major clusters represent broader groups, while smaller clusters indicate more specialized niche topics. The density map overlay (Figure 20) helps in further understanding the clusters colliding with each other.









Francisco at the top universities in terms of citation numbers as well. This university has 14241 citations at the time of writing this paper which shows the fact that this organization has had an increase in the number of publications in addition to maintaining the quality of the articles. In the second place, we have Harvard University with 12075 citations followed by the University of Toronto with 11452 citations.

Countries' citation and publication patterns were not as balanced as the other two. The USA topped the charts with 4312 articles containing 151249 all-time citations. The UK followed second in both tables with 1476 articles related to Medical English generating 37535 citations as of 2nd January 2024. Following the trend, Canada was placed third in both charts with 817 articles gaining 29625 citations overall.

Moving to the most important part of scientometric studies, theme, and keyword analysis was done by addressing the two types of keywords existing in the dataset. Care, impact, health-care, health, and quality were most prominent across the documents. In addition to the keyword and theme analysis, the co-occurrence of the keywords was analyzed to gain an in-depth snapshot of the relations and connections between the words

These findings can help scholars, stakeholders, and educators find their way in the intertwined academic field of Medical English. Understanding the trends in Medical English can help future training programs and guide research priorities. Researchers can also use these findings and predict the future of Medical English as a discipline, find the trending topics, and in the end, make this discipline trending once more.

As for the limitations of this study, although we have thoroughly analyzed every document accessible to us, some of the articles, including the top documents based on citation patterns, were not related directly to Medical English. This happened because the authors of those articles have written “Medical English” as one of the keywords of those documents. A more precise analysis containing fewer documents but more in-depth analysis could reveal a more error-free result in terms of top documents cited in the field.

Another avenue for researchers to explore can be the analysis of abstracts within the literature. This can help us achieve total control over the thematic evolution of the field. This is because the usual analysis of themes is done by just titles, reference titles, and the actual keywords. However, if there is a thematic analysis of abstracts within the literature, we can precisely plan future directions for the academic programs. Additionally, scholars can address these techniques for each of the three clusters that were mentioned before to gain a more precise map of each domain. We hope to see more scientometric and bibliometric analyses like this study for other disciplines to gain a better understanding of what has happened and what we can expect to come next.

## References

- Bachman, L. F., Palmer, A. S. (1996). *Language testing in practice*. Oxford University Press.
- Belcher, D. (2009). What ESP is and can be: An introduction. In D. Belcher (ed.), *English for specific purposes in theory and practice* (1–20). University of Michigan.
- Benesch, S. (2001). *Critical English for academic purposes*. Lawrence Erlbaum
- Dudley-Evans, T., St John, M. (1998). *Developments in English for specific purposes: A multidisciplinary approach*. Cambridge University Press.
- Hutchinson, T. & Waters, A. (1987). *English for specific purposes: A learning-centered approach*. Cambridge University Press.
- Hyland, K. (2006). *English for academic purposes*. Routledge.
- Kessler, M. (1963). An experimental study of bibliographic coupling between technical papers. *IEEE Transactions on Information Theory*, 9(1), 49–51. <https://doi.org/10.1109/tit.1963.1057800>
- Long, M. H. (2005). *Second language needs analysis*. Cambridge University Press.
- Paltridge, B., & Starfield, S. (Eds.). (2013). *The handbook of English for specific purposes*. John Wiley & Sons.
- Richards, J. C. (2001). *Curriculum development in language teaching*. Cambridge University Press.

- Salager-Meyer, F. (2014). Origin and development of English for medical purposes. Part I: Research on written medical discourse. *Medical Writing*, 23 (1), 49-51. <https://doi.org/10.1179/2047480613Z.000000000187>
- Ullah, M., Shahid, A., Din, I. ud, Roman, M., Assam, M., Fayaz, M., Ghadi, Y., & Aljuaid, H. (2022). Analyzing interdisciplinary research using co-authorship networks. *Complexity*, 2022, 1–13. <https://doi.org/10.1155/2022/2524491>
- Van Eck, N., & Waltman, L. (2010). Software survey: VOSviewer, a computer program for bibliometric mapping. *Scientometrics*, 84(2), 523-538. <https://doi.org/10.1007/s11192-009-0146-3>
- Zhang, J., Yu, Q., Zheng, F., Long, C., Lu, Z., & Duan, Z. (2015). Comparing keywords plus of WOS and author keywords: A case study of patient adherence research. *Journal of the Association for Information Science and Technology*, 67(4), 967–972. <https://doi.org/10.1002/asi.23437>
- Zhou, X., Zhou, M., Huang, D., & Cui, L. (2022). A probabilistic model for co-occurrence analysis in bibliometrics. *Journal of Biomedical Informatics*, 128, 104047. <https://doi.org/10.1016/j.jbi.2022>.



## Mapping the Antecedents of Foreign Language Enjoyment in Online Education: A Qualitative Diary Analysis

Fateme Chahkandi<sup>1\*</sup> 

<sup>1</sup>Assistant Professor in Applied Linguistics, Department of English Language, University of Birjand, Birjand, Iran

### ARTICLE HISTORY

Received: 30 April 2023  
 Revised: 01 June 2023  
 Accepted: 25 August 2023  
 Published: 30 September 2023

### CORRESPONDING AUTHOR

E-mail: [f.chahkandi@birjand.ac.ir](mailto:f.chahkandi@birjand.ac.ir)

### ABSTRACT

The present study set out to chart the experience of Foreign Language Enjoyment (FLE) and its antecedents in online education settings using the genealogy of emotion (Zembylas, 2002, 2008) as the theoretical framework. To fulfill this aim, e-learners in EFL classes were asked to keep emotion diaries for a semester-long period. The emotion diaries reporting the experience of FLE were then analyzed using Pekrun's (2014) taxonomy of emotions in connection to FLE including achievement-related FLE, epistemically-related FLE, topic-related FLE, and socially-related FLE. Analysis of results indicated that teachers, peers, and family members were the main sources of FLE in the social category. Additionally, achievement in exams and classroom tasks and activities, topics related to students' local cultural and social context, in conjunction with cognitively demanding tasks and questions that required learners to struggle over finding the answers were reported as other antecedents of enjoyment in online setting. It was discussed that Pekrun's model (2014) is a good fit for the analysis of the origin of single emotions like FLE as the previous data-driven analyses missed out on some of the main antecedents of FLE. Finally, the theoretical and practical implications were discussed especially for teachers.

**KEYWORDS:** Academic emotions; Emotion diary; Foreign language enjoyment; Online learning

## 1. Introduction

Enjoyment is defined as a sense of novelty or accomplishment experienced when individuals are capable of going beyond themselves to achieve something unexpected and is characterized by a high degree of engagement in an activity without noticing time and self (Csikszentmihalyi, 2008). It pushes participants into action and hones their motivational level against upcoming difficulties (Dewaele & MacIntyre, 2016). Foreign language enjoyment (FLE) is also related to positive emotions that learners experience when they are engaged in learning or using a foreign language (Dewaele & MacIntyre, 2014; Dewaele et al., 2018). It positively affects aspects of second language acquisition including learners' motivation, willingness to communicate, and their ultimate learning achievements (Botes et al., 2022; Dewaele & Dewaele, 2018).

FLE as one of the most frequent and salient emotions experienced by learners (Jiang & Dewaele, 2019; Piniel & Albert, 2018, Zhang, et al., 2021) has been given due attention in brick-and-mortar class contexts (e.g. Dewaele, 2022; Dewaele & Alfawzan, 2018; Li, et al., 2018). However, its investigation in online EFL learning settings warrants more attention. Yet, the issue of FLE in remote education is of particular significance given that students may experience lower levels of FLE as a result of technological failure and lack of familiarity with different tools and resources (Bond & Bedenlier, 2019). Moreover, being physically distant from their peers and teachers, e-learners may experience more negative emotions than their

counterparts in in-person classroom settings. In the face of these difficulties in virtual learning, positive emotions and enjoyment are necessary components to keep participants on track (Zhang et al., 2021). In addition, FLE positively affects participants' engagement in online learning (Deng et al., 2022; Derakhshan & Fathi, 2023; Guo, 2021; Luo et al., 2023). Experiencing positive emotion of enjoyment can boost participants' motivation to participate in activities, have interactions with peers and instructors, and show persistence in their language learning journey (Derakhshan & Fathi, 2023). In effect, FLE in cyber-based learning environments may counteract some of the negative effects arising from isolation and lack of interpersonal interaction and communication (Simamora, 2020).

The aim of the present study is to delve into EFL students' experience of FLE in online courses. Attention to the role of emotions in e-learning contexts is required to accomplish the goal of developing positive attitudes toward online education. Furthermore, as the interest in web-based pedagogical environments is likely to increase, such an analysis is crucial for creating learning environments which lead to higher student learning and performance. It will also provide faculty with greater awareness of the emotional components of web-based education in addition to corrective practices and measures they need to take in order to instigate positive emotional experiences and minimize the negative feelings associated with online learning.

## 2. Literature Review

### 2.1. Research on foreign language enjoyment

Research on the role of emotions in second language acquisition (SLA) has gained momentum over the past years. Yet, emotions have remained an under-researched area considering that the cognitive processes have dominated the field of SLA for decades (Dörnyei & Ryan, 2015). In fact, as Shao et al. (2019) asserted research on the role of emotions in SLA is still lagging behind the rapid embracement of emotions in psychology and education. Additionally, it has been primarily dominated by research on negative emotions and in particular foreign language anxiety (Daubney, et al., 2017). Nonetheless, as Dörnyei and Ryan (2015) put it, it's time to address the 'emotional deficit' in SLA research and to accommodate more positive emotions into our descriptions of learner psychology.

Previous literature demonstrates that positive emotions comprising enjoyment of learning, hope, pride, and relief have been reported with the same frequency (Pekrun, et al., 2002). By and large, positive feelings are described in relation to feelings of success, lack of anxiety, and enjoyment linked to academic accomplishments (Piniel & Albert, 2018). Enjoyment, in particular, has been linked to lower levels of anxiety, higher relative achievement position in the class, learner-internal variables such as proficiency levels and attitudes, and more foreign language use by the teacher (Dewaele et al., 2018). It is also related to the frequency of target language use and students' perception of self (Saito et al., 2018).

Zhang et al. (2021) identified three broad research strands in the literature on FLE in the context of SLA. First are studies which aim to measure FLE. Attempts in this regard include the development of the Foreign Language Enjoyment scale (Dewaele & MacIntyre, 2014) encompassing enjoyment in relation to the learners' subjective experience, their teachers, and peers. Extensions of the same scale have been carried out by Dewaele and MacIntyre (2016), Dewaele and Alfawzan (2018), and Dewaele, et al. (2019). In addition to the scales, idiosyncratic studies have been conducted which aim to depict the dynamics of learners' experience and their trajectories using multiple instruments including interviews, questionnaires, and journals. Examples of research in this category include Elahi Shirvan, et al. (2020); Dewaele and MacIntyre (2014), and Jiang and Dewaele (2019).

The second strand of research concerns studies which seek to identify the effect of different teacher-related or learner-related variables on FLE. In particular, FLE has been reported to be mainly instigated by teacher-related factors including teachers' friendliness, use of jokes, their frequency of English use (Dewaele & MacIntyre, 2014; Dewaele, et al., 2019; Jiang & Dewaele, 2019). Teachers' personality traits including openness, extraversion, and agreeableness impacts students' FLE as well (Ahmadi-Azad, et al., 2020). With regard to students' variables, the literature also indicates that there is a positive correlation between learners' grit and FLE (Yang, 2021) as well as between ideal L2 self and FLE (Liu, et al., 2022). Learners' gender is also related to FLE with female students experiencing higher levels of FLE than their male counterparts (Liu, et al., 2022). Additionally, students with more favorable attitudes toward English and English teachers tend to experience higher levels of enjoyment (Dewaele, 2022).

The third line of research centers on the relationship between FLE and learners' performance. FLE has been shown to improve learners' ability to notice input and hence increases the likelihood of absorbing foreign language. Moreover, it makes learners resilient and hard in the face of difficulties (Pavelescu & Petric, 2018; Prior, 2019). Results from Yuan (2023) also indicated that participants experiencing higher levels of enjoyment tend to experience lower anxiety levels. That being said, it needs to be noted that the existing literature mainly concerns traditional classrooms where the active participation of the teacher and peers is taken for granted. Nonetheless, as Regan et al. (2012) argue, the emotional tone of online environments differs from that of traditional educational conditions. Considering the significance of enjoyment in online education, this study aims to explore the sources of FLE in online foreign language classrooms.

## 2.2. Emotions in the context of online education

Technology-based delivery modes of education have been criticized for lacking some of the emotional features of on-campus teaching including facial expressions and body language (Vrasidas & Zembylas 2003). With students being distant from instructors and peers, having trouble interacting with them, and being obliged to undergo the demands of self-directed learning (Hawkins et al. 2012), they reportedly experience more negative emotions in internet-based learning settings (Stephan, et al., 2019). Yet, some studies do report on students' positive emotional experiences.

In a pioneer study on learners' emotions in the context of an online distance learning program, Zembylas, et al. (2008) discovered that most of the positive emotions were directly linked to the advantages of distance learning including excitement for the flexibility of distance learning, the emotional nature of distance communication, the possibility of developing social relationships with instructors and peers along with satisfaction for fulfilling the course requirements. Hilliard et. al.'s (2020) findings from research on the emotional experience of distance learners undertaking an online, collaborative group project also indicated that satisfaction and relief were the most frequently reported positive emotions. The sources of emotions were also mainly related to the completion of projects or modules, participation and engagement from others, task progress, working and collaborating with others, performance in the project or module, self-belief, the task itself, guidance and support.

Other factors contributing to positive emotions in e-learning environments include higher levels of perceived achievement and learner satisfaction, systematically structured and motivational instructional content which is aligned with learning objectives, adoption of a variety of appropriate evaluation methods, and homework and class participation being considered a part of final grading (Lee & Chei, 2020). Effective teaching activities such as direct instruction, receiving feedback, and multimedia resources can also promote positive emotions in e-learners (Yates, 2020).

Considering the focus of this study, FLE in the context of online education, the research identifies 4 sources for students' FLE (Wang, et al., 2021). The sources include teachers' traits including their praise and verbal encouragement, as well as online interaction, private learning performance in the form of personal progress, and students' academic achievement. In a similar vein, Yuan (2023) classified sources of FLE in online settings as self-, teacher-, peer-, and internet-related factors. Examples of self-induced sources of enjoyment included answering the questions, understanding the teacher, finishing assignments, and achievements in class. Teacher-related factors included teachers' praise, encouragement, personality, way of teaching, and sharing experiences. Peers' cooperation, discussion, interactions, and good classroom atmosphere were sources of enjoyment from peers. Finally, ways of online learning, online tests, and the Internet stability were mentioned with reference to the Internet.

Using the Q methodology, Thumvichit (2022) explored EFL learners' subjective experience with regard to FLE. Using principal component analysis and varimax rotation, the author extracted common patterns from participants' responses to a set of statements. The results of the analysis pointed to 3 distinct viewpoints. The first focuses on teachers' characteristics and practices where teachers contribute meaningfully to students' learning through being accessible and friendly and providing them with opportunities to share their ideas. The second viewpoint concerns students' autonomy which refers to students' ability to use the online resources to support their own learning. The third one relates to the learners' social experiences which emphasized students' receiving support from their peers and working as a team.

In addition to the sources of FLE referred to above, teachers' competency in internet and communication technology, (Shao, et al, 2023), the amount of teacher and peer support available, students' satisfaction with and personal pride in the classroom, learning atmosphere, and their personal feelings (Apridayani & Waluyo, 2022) have been reported to affect students' experience of FLE.

## 2.3. Theoretical framework

This research draws on the notion of genealogy of emotions (Zembylas, 2002, 2008). This method provides a way of understanding emotions through examining how learners' emotional experiences are shaped and reformed during the process of learning in a particular social context. In fact, genealogies of emotion cast light on the way events, objects, and persons contribute to the realization of emotions. This implies that establishing genealogies of emotions entails looking into the historicity of emotions and tracing their origins and antecedents. Regarding the genealogies of FLE in online learning, it focuses on the analysis of learners' experiences of enjoyment associated with online learning, the historicity of FLE and the way it unfolds and changes over a semester-long period, as well as the factors giving rise to it as its antecedents.

This study is also informed by Pekrun's (2014) classification of academic emotions. This framework provides a taxonomy of emotions related to academic contexts including achievement emotions, epistemic emotions, topic emotions, and social emotions (Pekrun, 2014). Achievement emotions are associated with academic activities and success or failure pertinent to these activities. Examples include enjoyment of learning; hope and pride connected to success; and anxiety and shame associated with failure. Epistemic emotions concern cognitive challenges such as surprise about a new task; curiosity, confusion

and frustration at facing barriers; and delight at solving the problems. Topic emotions are emotions provoked by topics presented in lessons. They can be both positive and negative and can constitute empathy with the characters of the story, anxiety and disgust resulting from dealing with medical issues or enjoying a painting in an art class. Lastly, social emotions are linked to interaction with teachers and peers as well as group learning in the classroom. Examples are love, sympathy, compassion, admiration, contempt, envy, anger or social anxiety. In the present study attempt is made to examine how enjoyment as a single academic emotion can be linked to each situation, i.e. achievement, epistemic challenges, topics, and social situations. It is believed that in comparison to the data-driven approach taken by prior research (e.g. Wang, Yang, and Wang, 2021; Yuan, 2023) this framework presents a more comprehensive and complete picture of the FLE in online settings. Accordingly, the following research question is proposed and investigated throughout the current study:

What are the antecedents of FLE in online education settings?

### 3. Methodology

#### 3.1. Setting and participants

This study was conducted in the context of a state university in Iran during the academic year 2021-2022 in a series of online EFL courses. The courses in which the study was carried out were mainly English courses offered to the EFL students including English grammar, Storytelling, Study Skills, Reading, and Speaking and Listening. At the beginning of the academic year, volunteered undergraduate students were recruited for this study. The purpose of the study was explained to them and they were asked to keep emotion diaries for every session of the courses they had taken and send them back to the researcher as the sessions ended. A total of 87 undergraduate students in English Literature and Translation Studies agreed to take part in the research project. They were 63 female and 24 male students with an average age of 24 (19- 29 years old). They were in the first or second year of remote learning and the platform for delivering the online courses was Adobe Connect where the course content was delivered synchronously. The participating students received credits for their cooperation in the study.

#### 3.2. Instrument and procedure

Emotion diaries were selected as the means of data collection in the present study since they compensate for data loss, reduce memory bias (Wallbott & Scherer, 1989), and allow for the exploration of emotions and their origins at multiple points of time along with temporal variations and within-person changes (Becker, et al., 2015; Hilliard et al., 2020). This method was considered to be a good fit for constructing the genealogy of emotions. The emotion diary was adapted from Averill (1983), Oatley and Duncan (1994), and Zembylas (2002) and was developed in five sections.

The first question asked participants' demographic information including age and sex. The second question asked students about the types of emotion(s) they experienced during the episode of the online course. To do so, they could select from among 10 positive emotions including joy, enthusiasm, satisfaction, pride, love and affection, intimacy, relief, empathy, enjoyment, and comfort along with 11 negative emotions comprising disappointment, disgust, guilt, boredom, anger, anxiety, surprise, fear, sorrow, sadness, and powerlessness. The list of emotions was adapted from Pekrun, et al. (2002) with some emotions being added during the pilot testing of the questionnaire. Following, they specified those who caused the emotion including the respondent himself/herself, the instructor, peers and classmates, parents, or others. In the next questions, they noted the incident (what happened?) and the antecedents of emotion (what made you feel this way?) in an open-ended format.

The diary questionnaire was translated into Persian and pilot-tested with 10 students. Participants were required to record every event that was emotionally salient instantly after or during the online class and send it back to the instructor after the class for a whole educational semester. However, only those diaries which reported experiencing enjoyment were considered for the current study. The total number of diaries received over the course of one semester included a corpus of 528 diary entries out of which 126 were on enjoyment.

#### 3.3. Data analysis

Content analysis was utilized to analyze the data obtained from the diary entries. First, the diary journals were read and summarized for individuals' reference to FLE and the incidents that caused the emotion. The summaries were coded using Pekrun's (2014) taxonomy of emotions in connection to FLE including achievement-related FLE, epistemically-related FLE, topic-related FLE, and socially-related FLE. Initially, the entries were read to familiarize ourselves with the data and notes of summaries were taken to get the initial ideas for coding. Next, we looked for the closest taxonomies which fit the initial codes. The coding phases were performed by two researchers working independently. Once the data were fit into taxonomies, the analysis was double-checked and the contradictory points were negotiated between the 2 raters until they reached the agreed-



upon analysis. further re-readings were done to ensure that all pieces of data had been incorporated into the analysis. Finally, segments of verbatim quotes were selected to illustrate the emergent themes.

## 4. Results

The results of diary analysis indicated that students experienced FLE in online settings mainly in relation to social dynamics and classroom atmosphere (74 tokens), followed by FLE as a result of achievements (63 instantiations), and topics raised in the EFL classes (33 examples). The least FLE also belonged to epistemic cognitive activities (28 instances).

### 4.1. Socially-related FLE

The sources of FLE in this category primarily came from teachers, peers, and family members respectively.

#### 4.1.1. Teachers as sources of FLE

Most of the incidents of FLE were experienced by e-learners in connection to the social relations in which they were involved. First and foremost, the teachers acted as the main source of students' FLE. This emotion was specifically provoked by the teacher's quality of teaching including clear and easy-to-understand explanations, content mastery, pedagogical and technological knowledge and skills in conjunction with the use of technological tools and applications. One of the participants illustrated how he enjoyed the teacher's teaching practice along these lines. *"The teacher explains the lesson in a way that everyone can comprehend. Also, she plays videos for us which are extremely engaging and help substantiate the structures in our minds. I enjoy her teaching and the use of up-date resources"*.

In addition, FLE was linked to teachers' rapport-building behavior including friendly acceptance of students, showing respect for them, seeking and welcoming their perspectives, perceiving their problems and striving to remedy those problems, and providing opportunities for students to compensate for their poor performance, lacks and inadequacies.

At the start of the semester, the instructor had a chit-chat with the students which created a good relationship with us. We thought she was by no means far away although we were in an online class and we felt she could perceive our struggles... The teacher is never conceited and always has modest behavior.

E-learners valued teachers' humane and friendly behavior in increasing their enjoyment where the distance between them and teachers was minimized by teachers' welcoming and approachable manner and emphasized that in a stress-free classroom environment, they did not need to worry about the score or teacher's reprimand allowing them to give their full attention to the class and enjoy learning new things.

In the speaking course, the class atmosphere is relaxed and intimate and we can talk without worrying about scores or anything. I don't feel humiliated in that class. The instructor and students are all welcoming and it gives me the feeling that I can succeed in this major.

Comments were also expressed in relation to the mood in which the teacher started the class where the same mood was carried over to the students during the class. Furthermore, students reported experiencing enjoyment when the teacher gave them hope for their future careers and professional lives.

The instructor discussed the prospects of this major with us. I enjoy it when teachers make connections with our future. It pushes me to think about my future more seriously and boosts me to continue with more determination and strength. It also drives me to keep working hard to accomplish my goals.

Teachers' humor and personalized stories which were intended to change the class tone and humanize the remote impersonal classes were also reported as additional sources of FLE. Moreover, FLE was enhanced when teachers gave recognition of students' progress or acknowledged their collaboration and active participation in class. *"In class, the teacher noticed my correct response and thanked me for that. I take pleasure when the instructor is happy with me"*. Accordingly, teachers' expressing their satisfaction with students and complimenting their efforts were reported to instigate students' FLE. *"I was doing the grammar exercises and I felt improvement in myself. In fact, the teacher's compliments showed this. It promoted my participation in class and often I volunteered to do the activities with no fear and stress"*.

#### 4.1.2. Peers as sources of FLE

In addition to teachers, peers and classmates were also found to be sources of FLE in cases where they provided support and encouragement, especially in the face of problems and difficulties. This was of particular significance to the participants considering that while they were learning remotely from each other, the intimacy and warmth of their support could be felt

beyond the remote distance. Peers' positive comments on each other's performance, explanations, and sharing points as well as the stories they shared were of particular importance to e-learners' experience of FLE. Also, their collaboration in class activities and the resultant constructive competitiveness made the class atmosphere lively and vibrant which gave rise to FLE on the part of the students.

The warmth and strength that my classmates give me motivate me to keep going. I'd like to save this energy so that I could have the power forever. I feel positive about their collaboration in class. The intimate atmosphere gives me hope and relief. This friendship is a turning point for me. I like their company.

#### 4.1.3. Family as a source of FLE

Given that the virtual classes are held at home with family members' presence, they can be of great help or on the other hand, a hindrance to online learners. As a source of FLE, the family was referred to in cases where they were perceptive of students' conditions by keeping quiet while they were involved in online classes and exams.

I was taking my exam online and all the family was quiet without a single word. I took my exam in peace and quiet and fortunately, it was a success. I'm glad my family values me and perceives my condition. It gives me great pleasure to be with them.

Another instance included a family member accompanying students while in class. Additionally, family's encouragement, respect, care, and pride in students' achievements and belief in their abilities were found to be associated with the experience of enjoyment.

I was in the class when my dad came into the room and accompanied me for the class. The importance and value that my dad gives to me and my affairs really matter to me. I feel delighted when my family values my education and my field of study.

## 4.2. Achievement-related FLE

Enjoyment was experienced as a result of students' accomplishments in EFL classrooms either on particular occasions including their good performance on tests and exams, giving the right answers to the questions posed by the teacher, and having a good presentation, or in the overall language learning process.

It was my first quiz online. I was totally stressed out. I was really worried about running out of time or losing internet access. I rushed through the exam and speeded it up. Luckily, I was able to make it. I was delighted.

Some participants reported enjoyment on the grounds that they felt progress and improvement in learning English and were able to gain mastery of diverse aspects of language such as speaking, pronunciation, reading, and grammar. More instances of achievement included the ability to concentrate, show diligence, and overcome anxiety and lack of confidence as the pre-requisite skills for learning.

It was my first time giving the summary of "TED talks" in class. I was stressed out. To be honest, I'm afraid of my own voice being heard over the mic. I turned on the mic and started giving my summary. At that moment, I was just thinking that I could. I thought to myself, if I wanted to overcome stress, I had to start from somewhere. It went all well and my summary was done. It wasn't as hard as I had thought.

Understanding the lessons especially the ones being formerly hard to understand or the ones which included high levels of English proficiency such as stories and texts also stimulated students' enjoyment.

The instructor said the lesson would not be easy and we needed to pay careful attention. However, I learned it while my classmates were still struggling with it. I realized that I would learn if I were attentive enough. I should just believe in myself.

Some also contended that they could have a relative standing among their peers by providing the right answer or volunteering to do the assigned tasks and exercises. *"The instructor asked a question the answer to which nobody knew. When I was able to answer the question, I felt proud of myself"*. Along the same line, this quote from a different student illustrates how enjoyment comes from taking on challenging tasks. *"When the instructor checks the homework in class and reminds us of our issues, I feel double enjoyment as I learn the point even deeply"*. Additionally, learning while multitasking in online classes was mentioned as an instance of achievement by one of the students. *"I had some guests to come over and I was preparing dinner while I was listening to the class. Although I was multitasking, I managed to learn the lesson."*

## 4.3. Topic-related FLE

This category of FLE was associated with the topics raised in online classes or general interest in English as a subject for learning. First, a multitude of students reported enjoyment coming from the local stories presented by their peers in the storytelling class. These stories were regarded as ties to the students' own culture and history through which they got familiar with distinguished people, locations, traditions, costumes, and legends of their country. Also, they alleviated students' hardships in life through both instigating positive feelings and making students' problems appear insignificant to them.

With the many problems I have in my life and the overwhelming despair and anxiety stretching over my life, I receive full enjoyment of the local stories told in class. when I listen to the stories of the past people, I ponder over how they could survive all that hardship. Then, I feel less disappointed, and, although momentarily, I feel myself immersed in the story. I try to remind myself of the story during the day to make my problems appear small.

Moreover, the stories were believed to raise students' interest and curiosity and were a drive for them to keep track of the class up to the end." *I had a good feeling. I felt proud of my country and the people. The topics arouse my curiosity and made me follow the class. Although the classes are online, I don't feel bored or digressed*". By the same token, students' narratives recounted in response to grammar questions and the opportunities for personalizing the lessons impacted their learning and the recall of the lessons and were considered as a further source of enjoyment in EFL classes. "*When my friends were talking about their past habits and states using the structure of "used to", we were amazed by how strange they were. It was all fun.*"

In addition to the situation-specific topics pinpointed above, English was also viewed as an interesting subject of study by some students. Accordingly, students expressed FLE in relation to English movies, learning about foreign cultures and countries, as well as English literature and people's lifestyle, "*My interest in English causes me to attend the class with great enthusiasm and enjoy every moment of it. ...I suppose my interest makes the lessons more understandable to me*".

#### 4.4. Epistemically-related FLE

Epistemically-related FLE included enjoyment experienced as a result of some cognitive challenges involved in solving a problem, understanding something, or arriving at the answer to a question. Although this category of enjoyment was experienced less frequently than the previous ones, the results obtained from diary analysis pointed to some instances where students had to struggle cognitively to understand a grammatical structure, answer the challenging questions raised by the teacher, understand a story by focusing on the main words, and make sense of the new lessons. One example includes a diary entry in which the student described the struggle she had with learning a structure.

To me, the conditionals have been always tricky. I didn't learn them at school. However, today, I managed to learn them completely. I listened attentively and took notes. I also concentrated on details. I owe this to my instructor's clear explanations and the enthusiasm she gives us to follow the class."

Often, students described the strategies they employed to get over the cognitively demanding situations. In one case, a student described how he managed to employ advance organizers to activate his background knowledge in order to comprehend the new lesson. "*I decided to pre-study the lessons. This relieved the burden and helped me understand the lesson better. I also boosted my confidence, so I could listen with more relief and concentration as the lesson was not unfamiliar anymore*". Another student noted that she watched the class archive to further reinforce her learning and when she noticed her progress and preparedness in class, she felt extreme enjoyment. Still, a further student talked about how he increased his concentration. "*I decided to keep track of the class and stay tuned. when I see myself part of the class and increase my participation, I feel I can pass the course with a better score*". Additionally, the cases where teachers challenged students' knowledge by raising demanding questions and involving them in competitive activities for bonus scores were referred to as one of the most frequent antecedents of FLE.

We were doing the grammar exercises in class when the teacher raised a question and said she would consider a bonus score for the person who answered the question correctly. I had a strange feeling. Although English has been part of my life for a couple of years, I noticed that there were still a lot of things I didn't know.

Although the instances of FLE reported in this category referred to enjoyment induced by being involved in some cognitive challenges, students reported on another set of antecedents which pointed to students' psychological barriers and challenges being resolved and consequently, freed their mental capacity for understanding the lessons. Examples of such incidents include students' stress being over when they had their presentations in class or after finishing their exams. "*I had my presentation the other week and the relief coming from this made me drown in the new lesson. It even helped me understand the lesson completely as my full concentration was on the new lesson*".

## 5. Discussion

This article drew on Pekrun's (2014) model of academic emotions to delineate FLE in online classrooms. The findings obtained from content analysis of students' diaries over a semester-long period indicated that the model can be considered a good fit for portraying the dynamics of not only academic emotions but also highlighting the antecedents of single emotions like FLE. It was demonstrated that teachers' attributes had the most impact on students' experience of FLE including their pedagogical, content, and technological knowledge, their rapport-building behavior, as well as their recognition and acknowledgment of students' progress. Peers' interaction, encouragement, and scaffolding along with parents' respect, care, and motivation were also found to induce FLE in remote education. Other sources of FLE included achievements on exams and classroom tasks and activities, local stories of students' place of residence which bound students to their cultures and traditions, as well as the personalized topics recounted by teachers and students alike which provided for personalization of the content and materials. Finally, cognitively demanding tasks which required students to strive to solve a problem or come up with the answer to a question were also reported to stimulate students' FLE.

The results of this analysis are in line with prior literature on face-to-face learning environments (Dewaele & MacIntyre 2019; Dewaele, et al. 2019; Jiang & Dewaele 2019; Li, et al., 2018) where teachers were found to meaningfully contribute to students' FLE if they are accessible and friendly and if they encourage students to ask questions and to speak without worrying about correct structure. Likewise, teacher friendliness has been found to have a major role in students' FLE as was predicted by Dewaele et al. (2019). It was shown that as a result of teachers creating a friendly and non-threatening environment, students are more open to target language input and their capacity to absorb the input and to concentrate increases (Dewaele, 2011). In the Iranian context, too, teachers' personality traits such as openness and agreeableness were found to positively affect students' FLE (Ahmadi-Azad, et al., 2020)

In online educational contexts, similar to Yuan's (2023) findings, teacher's characteristics such as familiarity with the teaching platform and having mastery over it, friendliness, having online interactions with learners, joking, praise and encouragement, and their good personalities were found to positively impact students' FLE. In addition, teachers' use of relevant and interesting content such as videos created enjoyable learning experiences for e-learners. Accordingly, it highlights the significance as well as the difficulty of teachers' role in remote education. In other words, they bear more responsibility to create a welcoming atmosphere in the absence of face-to-face interactions and in-person social connections which makes students more prone to digression and lack of attention.

Similarly, the role of students' peers and the enjoyment taken from them was highlighted by our results. As was demonstrated by Li, et al. (2018), peers' engagement and collaboration can have a facilitative role in creating a warm classroom climate. The results in this part are also in line with those of Yuan (2023) where peers' collaboration to do the assignments, discussions on the lessons, interactions, and the good classroom atmosphere were reported to give rise to enjoyment. Positive emotions such as enjoyment are sociocultural artifacts that are constructed and regulated through social and interpersonal interactions (Gross & Barrett, 2011; Swain, 2013). In online collaborative learning environments, participants have been shown to exhibit high levels of enjoyment over the course of interactions with peers and FLE has been reported as an indispensable part of group work in online settings (Zhang, et al., 2021).

However, unlike the results in Thumvichit's study (2022), participants in this study provided no accounts of autonomous learning and the use of e-learning resources as a source of FLE in cyber-based learning. In a similar vein, and unlike Wang, et al.'s results (2021), our participants gave no account of the Internet interactions as a source of FLE. On the contrary, Internet connection was in most cases recounted as a source of negative emotions such as anxiety. The difference may reside in the access to reliable and stable Internet connections or students' lack of familiarity with or interest in the use of Internet-based resources. One further possible explanation may be that as students were asked to keep diaries of classroom events during which they felt particular emotions, their reports are bound to classroom contexts rather than out-of-class activities and events.

The study also highlighted the functions of emotions identified by MacIntyre and Gregersen (2012) including enhancing attention and cognition, mitigating the effects of negative emotions, encouraging resilience, in addition to developing social connections and interactions. The findings also revealed that, as stated by the control-value theory (Pekrun et al. 2007), FLE can be experienced both in aroused states where students are involved in challenging tasks and in relaxed states when they are relieved and can enjoy stress-free situations. However, prior studies on the antecedents of FLE rarely report on the role of topics or epistemic challenges in the experience of enjoyment. Likewise, they rarely note how parents and family members can affect the experience of FLE in online classes. More often than not, they relate FLE to teachers, peers, and self as the main sources. Yet, as was stated earlier, reliance on Pekrun's taxonomy (2014) enabled us to describe FLE in relation to a wider spectrum of situations and events including cognitively demanding tasks and situations alongside engaging topics.

## 6. Conclusion

This study demonstrated that fun and enjoyable experiences should be incorporated into any e-learning experience as it yields academic and psychological benefits and improves learners' cognitive abilities (Dewaele & MacIntyre, 2014). Theoretically,

the study offered some hints on how Pekrun's (2014) classification of academic emotions can be applied in understanding the multitude of conditions and experiences that shape a single emotion like FLE in online settings. Moreover, the employment of the genealogy of emotions (Zembylas, 2002, 2008) facilitated tracing the historicity of the emotion as it is tied to specific times and activities rather than providing an overall judgment of previously occurred activities and events as is often the case with the FLE scales. Practically, it provides the teachers with some guidance on the best practices and strategies that guarantee an enjoyable online learning experience. In particular, given that teachers constitute the main source of FLE even in the absence of face-to-face interactions, it is incumbent on teachers to adapt their pedagogical and personal characteristics to secure interconnectedness, and to optimize enjoyment in conditions where students are distant from each other. By the same token, changes can be made to ELT textbooks to ensure more enjoyable discussions and learning experiences.

Although illuminating, the study suffers from some limitations. The first limitation has to do with the nature of emotions themselves. As it was evident in the diaries, students' report of FLE was frequently mixed up with other positive emotions such as satisfaction, relief, happiness, hope, and pride. It seemed as though, participants themselves did not have clear-cut boundaries among the positive emotions in their minds. This makes finding the exact antecedents of FLE difficult, if not impossible. The second limitation concerns reliance on diaries as the only method of data collection. Future lines of research can back up diary analysis with video recordings of classroom events as well as interviews to obtain a deeper understanding of students' feelings. Also, further studies can be carried out on how other feelings unfold in the context of remote learning.

## 7. References

- Ahmadi-Azad, S., Asadollahfam, H., & Zoghi, M. (2020). Effects of teacher's personality traits on EFL learners' foreign language enjoyment. *System*, 95, 102369. <https://doi.org/10.1016/j.system.2020.102369>
- Apridayani, A., & Waluyo, B. (2022). Antecedents and effects of students' enjoyment and boredom in synchronous online English courses. *Journal of Multilingual and Multicultural Development*, 1-16. <https://doi.org/10.1080/01434632.2022.2152457>
- Averill, J. R. (1983). Studies on anger and aggression: Implications for theories of emotion. *American psychologist*, 38(11), 1145. <https://psycnet.apa.org/doi/10.1037/0003-066X.38.11.1145>
- Becker, E. S., Keller, M. M., Goetz, T., Frenzel, A. C., & Taxer, J. L. (2015). Antecedents of teachers' emotions in the classroom: an intraindividual approach. *Frontiers in psychology*, 6, 635. <https://doi.org/10.3389/fpsyg.2015.00635>
- Bond, M., & Bedenlier, S. (2019). Facilitating student engagement through educational technology: towards a conceptual framework. *Journal of Interactive Media in Education*, 2019(1). <https://eric.ed.gov/?id=EJ1228555>
- Botes, E., Dewaele, J. M., & Greiff, S. (2022). Taking stock: A meta-analysis of the effects of foreign language enjoyment. *Studies in Second Language Learning and Teaching*, 12(2), 205-232. <http://dx.doi.org/10.14746/ssllt.2022.12.2.3>
- Csikszentmihalyi, M. (2008). *Flow: The psychology of optimal experience*. New York: Harper Perennial.
- Daubney, M. A. R. K., Dewaele, J. M., & Gkonou, C. (2017). Preliminary thoughts on language anxiety and the focus of this anthology. *New Insights into Language Anxiety: Theory, Research and Educational Implications*, 1-7. <https://doi.org/10.21832/9781783097722-002>
- Deng, W., Lei, W., Guo, X., Li, X., Ge, W., & Hu, W. (2022). Effects of regulatory focus on online learning engagement of high school students: The mediating role of self-efficacy and academic emotions. *Journal of Computer Assisted Learning*, 38(3), 707-718. <https://doi.org/10.1111/jcal.12642>
- Derakhshan, A., & Fathi, J. (2023). Grit and foreign language enjoyment as predictors of EFL learners' online engagement: The mediating role of online learning self-efficacy. *The Asia-Pacific Education Researcher*, 1-11. <https://doi.org/10.1007/s40299-023-00745-x>
- Dewaele, J. M. (2011). Multilingualism: Acquisition and use. *International Journal of Applied Linguistics*, 21(2). [https://doi.org/10.1111/j.1473-4192.2011.00289\\_2.x](https://doi.org/10.1111/j.1473-4192.2011.00289_2.x)
- Dewaele, J. M. (2022). The interplay between learner-internal variables and levels of anxiety and enjoyment among Spanish EFL learners. In *Understanding variability in second language acquisition, bilingualism, and cognition* (pp. 217-239). Routledge.

- Dewaele, J. M., & Alfawzan, M. (2018). Does the effect of enjoyment outweigh that of anxiety in foreign language performance?. *Studies in second language learning and teaching*, 8(1), 21-45. <https://www.ceeol.com/search/article-detail?id=624879>
- Dewaele, J. M., & Dewaele, L. (2018). Learner-internal and learner-external predictors of willingness to communicate in the FL classroom. *Journal of the European Second Language Association*, 2(1), 24-37. <http://doi.org/10.22599/jesla.37>
- Dewaele, J. M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in second language learning and teaching*, 4(2), 237-274. <https://www.ceeol.com/search/article-detail?id=79515>
- Dewaele, J. M., & P. D. MacIntyre. (2016). Foreign Language Enjoyment and Foreign Language Classroom Anxiety: The Right and Left Feet of the Language Learner. In *Positive Psychology in SLA*, edited by P. D. MacIntyre, T. Gregersen, and S. Mercer, 215–236. Bristol: Multilingual Matters.
- Dewaele, J. M., & MacIntyre, P. D. (2019). The predictive power of multicultural personality traits, learner and teacher variables on foreign language enjoyment and anxiety. In *Evidence-based second language pedagogy* (pp. 263-286). Routledge.
- Dewaele, J. M., Magdalena, A. F., & Saito, K. (2019). The effect of perception of teacher characteristics on Spanish EFL learners' anxiety and enjoyment. *The Modern Language Journal*, 103(2), 412-427. <https://doi.org/10.1111/modl.12555>
- Dewaele, J. M., Witney, J., Saito, K., & Dewaele, L. (2018). Foreign language enjoyment and anxiety: The effect of teacher and learner variables. *Language teaching research*, 22(6), 676-697. <https://doi.org/10.1177/1362168817692161>
- Dornyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. Routledge.
- Elahi Shirvan, M., Taherian, T., & Yazdanmehr, E. (2021). Foreign language enjoyment: a longitudinal confirmatory factor analysis–curve of factors model. *Journal of Multilingual and Multicultural Development*, 1–19. <https://doi.org/10.1080/01434632.2021.1874392>
- Gross, J. J., & Barrett, L. F. (2011). Emotion generation and emotion regulation: One or two depends on your point of view. *Emotion Review*, 3(1), 8e16. <https://doi.org/10.1177/1754073910380974>
- Guo, Y. (2021). Exploring the dynamic interplay between foreign language enjoyment and learner engagement with regard to EFL achievement and absenteeism: A sequential mixed methods study. *Frontiers in Psychology*, 12, 766058. <https://doi.org/10.3389/fpsyg.2021.766058>
- Hawkins, A., Barbour, M. K., & Graham, C. R. (2012). Everybody is their own island: Teacher disconnection in a virtual school. *International Review of Research in Open and Distributed Learning*, 13(2), 123-144. <https://doi.org/10.19173/irrodl.v13i2.967>
- Hilliard, J., Kear, K., Donelan, H., & Heaney, C. (2020). Students' experiences of anxiety in an assessed, online, collaborative project. *Computers & Education*, 143, 103675. <https://doi.org/10.1016/j.compedu.2019.103675>
- Jiang, Y., & Dewaele, J. M. (2019). How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners?. *System*, 82, 13-25. <https://doi.org/10.1016/j.system.2019.02.017>
- Lee, J. Y., & Chei, M. J. (2020). Latent profile analysis of Korean undergraduates' academic emotions in e-learning environment. *Educational Technology Research and Development*, 68(3), 1521-1546. <https://doi.org/10.1007/s11423-019-09715-x>
- Li, C., Jiang, G., & Dewaele, J. M. (2018). Understanding Chinese high school students' foreign language enjoyment: validation of the Chinese version of the foreign language enjoyment scale. *System*, 76, 183-196. <https://doi.org/10.1016/j.system.2018.06.004>

- Liu, E., Wang, J., & Bai, S. (2022). Self-guides, enjoyment, gender, and achievement: a survey of Chinese EFL high school students. *Journal of Multilingual and Multicultural Development*, 1-18. <https://doi.org/10.1080/01434632.2022.2153854>
- Luo, R., Zhan, Q., & Lyu, C. (2023). Influence of instructor humor on learning engagement in the online learning environment. *Social Behavior and Personality: An International Journal*, 51(2), 1–12. <https://doi.org/10.2224/sbp.12145>
- MacIntyre, P., and Gregersen, T. (2012). Emotions that facilitate language learning: the positive-broadening power of the imagination. *Stud. Second Lang. Learn. Teach.* 2, 193–213. <https://www.ceeol.com/search/article-detail?id=72835>
- Oatley, K., & Duncan, E. (1994). The experience of emotions in everyday life. *Cognition & emotion*, 8(4), 369-381. <https://doi.org/10.1080/02699939408408947>
- Pavelescu, L. M., & Petrić, B. (2018). Love and enjoyment in context: Four case studies of adolescent EFL learners. *Studies in Second Language Learning and Teaching*, 8(1), 73-101. <https://www.ceeol.com/search/article-detail?id=624892>
- Pekrun, R. (2014). *Emotions and learning* (Vol. 24, pp. 1-31). Geneva, Switzerland: International Academy of Education (IAE).
- Pekrun, R., Frenzel, A. C., Goetz, T., & Perry, R. P. (2007). The control-value theory of achievement emotions: An integrative approach to emotions in education. *Emotion in Education*, 13–36. <https://doi.org/10.1016/B978-012372545-5/50003-4>
- Pekrun, R., Goetz, T., Titz, W., & Perry, R. P. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist*, 37(2), 91–105. [https://doi.org/10.1207/S15326985EP3702\\_4](https://doi.org/10.1207/S15326985EP3702_4)
- Piniel, K., & Albert, A. (2018). Advanced learners' foreign language-related emotions across the four skills. *Studies in Second Language Learning and Teaching*, 8(1), 127-147. <https://www.ceeol.com/search/article-detail?id=624895>
- Prior, M.T. (2019). Elephants in the room: An “affective turn,” or just feeling our way? *The Modern Language Journal*, 103(2), 516–527. <https://doi.org/10.1111/modl.12573>
- Regan, K., Evmenova, A., Baker, P., Jerome, M. K., Spencer, V., Lawson, H., & Werner, T. (2012). Experiences of instructors in online learning environments: Identifying and regulating emotions. *The Internet and Higher Education*, 15(3), 204-212. <https://doi.org/10.1016/j.iheduc.2011.12.001>
- Saito, K., Dewaele, J. M., Abe, M., & In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in classroom settings: A cross-sectional and longitudinal study. *Language Learning*, 68(3), 709e743. <https://doi.org/10.1111/lang.12297>
- Shao, K., Kutuk, G., Fryer, L. K., Nicholson, L. J., & Guo, J. (2023). Factors influencing Chinese undergraduate students' emotions in an online EFL learning context during the COVID pandemic. *Journal of Computer Assisted Learning*. <https://doi.org/10.1111/jcal.12791>
- Shao, K., Pekrun, R., & Nicholson, L. J. (2019). Emotions in classroom language learning: What can we learn from achievement emotion research?. *System*, 86, 102121. <https://doi.org/10.1016/j.system.2019.102121>
- Simamora, R. M. (2020). The Challenges of online learning during the COVID-19 pandemic: An essay analysis of performing arts education students. *Studies in Learning and Teaching*, 1(2), 86-103. <https://doi.org/10.46627/silet.v1i2.38>
- Stephan, M., Markus, S., & Gläser-Zikuda, M. (2019, October). Students' achievement emotions and online learning in teacher education. In *Frontiers in Education* (Vol. 4, p. 109). Frontiers Media SA. <https://doi.org/10.3389/educ.2019.00109>
- Swain, M. (2013). The inseparability of cognition and emotion in second language learning. *Language Teaching*, 46(2), 195e207. <https://doi.org/10.1017/S0261444811000486>
- Thumvichit, A. (2022). Unfolding the subjectivity of foreign language enjoyment in online classes: AQ methodology study. *Journal of Multilingual and Multicultural Development*, 1-18. <https://doi.org/10.1080/01434632.2022.2050917>




- Vrasidas, C., & Zembylas, M. (2003). The nature of technology-mediated interaction in globalized distance education. *International Journal of Training and Development*, 7(4), 271-286. <https://doi.org/10.1046/j.1360-3736.2003.00186.x>
- Wallbott, H. G., & Scherer, K. R. (1989). Assessing emotion by questionnaire. In *The measurement of emotions* (pp. 55-82). Academic Press. <https://doi.org/10.1016/B978-0-12-558704-4.50009-8>
- Wang, X., Wang, Y., Yang, Y., & Wang, L. (2021). Investigating Chinese university students' enjoyment in a web-based language learning environment: Validation of the online foreign language enjoyment scale. *Perceptual and Motor Skills*, 128(6), 2820–2848. <https://doi.org/10.1177/00315125211041714>
- Yang, B. (2021). Predicting EFL Learners' Achievement from Their Two Faces—FLE and FLCA. *Theory and Practice in Language Studies*, 11(3), 275-285. <https://doi.org/10.17507/tpls.1103.07>
- Yates, A., Starkey, L., Egerton, B., & Flueggen, F. (2021). High school students' experience of online learning during Covid-19: the influence of technology and pedagogy. *Technology, Pedagogy and Education*, 30(1), 59-73. <https://doi.org/10.1080/1475939X.2020.1854337>
- Yuan, R. (2023). Chinese university EFL learners' foreign language classroom anxiety and enjoyment in an online learning environment during the COVID-19 pandemic. *Asia Pacific Journal of Education*, 1-17. <https://doi.org/10.1080/02188791.2023.2165036>
- Zembylas, M. (2002). Constructing genealogies of teachers' emotions in science teaching. *Journal of Research in Science Teaching: The Official Journal of the National Association for Research in Science Teaching*, 39(1), 79-103. <https://doi.org/10.1002/tea.10010>
- Zembylas, M. (2008). Adult learners' emotions in online learning. *Distance Education*, 29(1), 71–87. <https://doi.org/10.1080/01587910802004852>
- Zembylas, M., Theodorou, M., & Pavlakis, A. (2008). The role of emotions in the experience of online learning: Challenges and opportunities. *Educational Media International*, 45(2), 107-117. <https://doi.org/10.1080/09523980802107237>
- Zhang, Z., Liu, T., & Lee, C. B. (2021). Language learners' enjoyment and emotion regulation in online collaborative learning. *System*, 98, 102478. <https://doi.org/10.1016/j.system.2021.102478>





<https://doi.org/10.22077/ali.2024.7257.1027>

## Critical Analysis of Complete IELTS Textbook: Cultural Factors in Focus

Somayeh Nourazar<sup>1</sup>   
 Mahdi Naseri<sup>2\*</sup>   
 Elham Mohammadi<sup>3</sup> 

<sup>1</sup>MA in Applied Linguistics for English Language Teaching, Department of English Language, Faculty of Humanities, University of Zanjan, Zanjan, Iran

<sup>2\*</sup>Assistant Professor in Applied Linguistics, Department of English Language, Faculty of Humanities, University of Zanjan, Zanjan, Iran

<sup>3</sup>Assistant Professor in Applied Linguistics, Department of English Language, Faculty of Humanities, University of Zanjan, Zanjan, Iran

### ARTICLE HISTORY

Received: 06 April 2023  
 Revised: 04 June 2023  
 Accepted: 11 August 2023  
 Published: 30 September 2023

### CORRESPONDING AUTHOR

E-mail: [mahdinasiri@znu.ac.ir](mailto:mahdinasiri@znu.ac.ir)

### ABSTRACT

The inseparable connection between language and culture has been at the forefront of many studies. Similarly, the role of teaching materials in fostering intercultural competence of language learners is consistently highlighted in the literature. In keeping with these observations, the current study aims at investigating the representation of culture in Complete IELTS textbook published by Cambridge University Press, 2012. Through adopting Kachru's concentric circles of English speakers around the world, the cultural content related to Center and Periphery Countries was explored. In order to quantify and categorize cultural themes, the Byram's cultural checklist (1993) was used. Next, the content of four sections of the textbook, namely reading, listening, speaking, and writing, was analyzed to record any cultural representation. The overall sample included 16 reading passages, 10 listening scripts, 10 speaking and 10 writing sections. The results indicated that Complete IELTS textbook contains cultural presentations of different countries with more weight given to Periphery Countries. The culture of these regions occupies 67.97% of the content in the textbook, whereas Center Countries take up 32%. Moreover, in terms of the eight areas of Byram, the majority of the cultural presentation (55.61%) falls under the category of geography. It can be concluded that the textbook is free of any cultural bias and ideological/political motives.

**KEYWORDS:** Textbook evaluation; Culture; Cultural content; Center and periphery countries

## 1. Introduction

Learning a language goes beyond linguistic knowledge and structure. In the process of learning a language, students are exposed to culture as it is embedded in the communication. Kramersch (1993) stated that language and culture are inseparable. He believes that EFL learners acquire both target language and culture. Cultural knowledge is part of the communicative competence which includes cultural understanding, conversational routines, and the target society's norms and values. Many other scholars believe in the strong relationship between language and culture and emphasize on the crucial role of culture in language learning and

teaching (Baker, 2012a; Baker, 2012b; Gao, 2006; Jamalvandi, 2013; Kramersch, 1993; Liddicoat and Scarino, 2013; Mitchell and Myles, 2004). Han and Bae (2005) stated that teaching a foreign language without its cultural content is impossible. Gao also (2006) emphasizes the close tie between language and culture stating that language learning is cultural learning and accordingly language teaching is cultural teaching.

Textbooks as an important educational tool in a language classroom serve an important role in teaching culture. They are sources of linguistic knowledge and cultural aspects as well and help learners gain sufficient cultural knowledge and awareness (Al-Sofi, 2018; McKay, 2002; Kramersch, 1993; Taki, 2008). Gray (2002) states that textbooks are designed to teach language, but they also carry cultural implications. Therefore, all textbooks are inclusive of cultural content and approaches in different ways.

The culture presented in textbooks should be fair. It is emphasized that English teaching and learning materials should encompass intercultural perspectives and present cultures from around the world rather than merely the culture of English-speaking countries (Baker, 2012; Bennett et al., 2003; Byram, 1997; McKay, 2002). As McKay (2002) puts, three cultures -the source, the target and international culture- should be included in English textbooks to fulfill the needs of different English learners.

Several studies have demonstrated that textbooks often contain cultural and linguistic biases, and social stereotypes (e.g. Amerian & Tajabadi, 2020; Ansary & Babaii, 2003; Dar & Masroor, 2019; Melliti, 2013; Shah & Pathan, 2016). A large body of research has also attested to the frequent use of target language culture elements in internationally published textbooks, and this, in turn, deprives learners of acquiring adequate knowledge for effective intercultural communication (e.g. Amerian and Tajabadi, 2020; Canagarajah, 2003; Chao, 2011; Gray, 2002; Melliti, 2013; Shah & Pathan, 2016). As an example, Amerian and Tajabadi (2020) have criticized the biased cultural representation and dominance of British culture in Oxford's New Headway (4th ed.) series.

Following the spread of English language in the world, Anglophone countries' industry, and specifically, Britain has widely benefited from the exportation of English textbooks (Gray, 2002; Melliti, 2013). Cambridge University Press (CUP) as the oldest university press in the world (Hill, 2016) publishes over 50,000 titles by authors from over 100 countries. In 2011, CUP formed a partnership with Cambridge Assessment to publish official preparation materials for Cambridge English and IELTS examinations.

The appropriacy of global textbooks is determined by their global use (Canagarajah, 1999; Gray, 2002). Therefore, internationally published textbooks (e.g. IELTS preparation textbooks) are worthy of in-depth investigation in terms of their users' perspectives and cultural content as well. Michael (1993 as cited in Melliti, 2013) refers to cultural content as one of the essential elements in analyzing the content of textbooks in addition to linguistic content.

The present study is an attempt to probe the cultural content of Complete IELTS textbook published by CUP (2012). The underlying theoretical framework adopted to this end is Kachru's famous concentric circles of English speakers around the world. According to his model, two major groups of countries are identifiable, i.e. Centre countries and Periphery countries. The former attributes to *inner* circle (English is used as a first language, e.g. UK, USA) and the latter includes both *outer* (English is used as a second language) circle and *expanding* (English is learned as a second language) circle.

In contrast to similar studies on internationally published textbooks, as mentioned earlier, claiming the significant representation of Center Countries' culture (Phillipson, 1992), the presentation of culture in the Complete IELTS textbook seems to be in favor of Periphery Countries, in particular UAE. The textbook seems to have instances representing local culture of UAE. The present study was conducted in Iran to address the numerous students and teachers' complaints of cultural bias. Upon further investigation, no previous study has been undertaken to assess cultural bias in the IELTS preparation textbooks, in particular Complete IELTS. Following Common European Framework of Reference (CEFR), CUP maintains fairness of its English language textbooks to all cultural groups. Therefore, with little evidence in the literature to support students and teachers' claiming cultural bias, a further investigation into the matter is needed to fill the wide gap in the literature. *Complete IELTS* textbook was investigated for the quantity and manner of presentation of Center and Periphery Countries' culture. This study aims at ensuring the neutrality of the textbook in terms of fair presentation of the world cultures. Moreover, this study provides feedback from a local context and makes textbook writers aware of some of the issues.

## 2. Literature Review

Since English is used as an international language, the cultural content of ELT materials should not be limited to native English-speaking cultures. According to Gray (2002) diversity of book users from around the world has to be embraced in designing a suitable global textbook.

Phillipson (1992, p. 17) classifies English language users around the world into two groups: Centre countries and Periphery countries. The term "Centre countries" covers Britain, the USA, New Zealand, Canada, and Australia. The majority of the population speaks English as their mother tongue. Countries outside this circle are called the Periphery countries. The

Centre countries are assumed to impose their culture to internationally published ELT textbooks. His argument revolves around the Centre countries' penchant for maintaining and extending their global influence through the English language teaching industry (Phillipson, 1992).

Some scholars in the field of ELT claim that the use of internationally published textbooks mostly produced by Western commercial organizations is not problem-free and may affect local curriculum; moreover, these textbooks are often considered as culturally inappropriate (Canagarajah, 2003; Gray, 2002). In a study on English language teachers' beliefs about internationally published materials in Indonesia, Zacharias (2005) mentions "teachers found some cultural aspects too remote to be understood or even the language too difficult" (p.33). It is not far from expectation that some learners may not have a clear understanding of some cultural aspects or they may find them inappropriate in their own culture.

This problem is not limited to global ELT textbooks. Some studies have proved the cultural bias in International English Language Testing System (IELTS) as well. In a study conducted in Bangladesh by Khan (2006), the subtle cultural biases toward Western culture and norms of behavior in the speaking test content have been revealed. Some of the topics on the test were identified as unusual, uncommon and unfamiliar. It was said that some vocabulary items and topics reflect Western concepts and patterns of interaction and are not culturally appropriate for local candidates. This makes the task difficult for them and affects their performance on the test. This issue needs test designers' attention.

In order to avoid such problems in ELT textbooks, publishers try to follow a "one-size-fits all" approach in designing global textbooks (Alkhatheery, 2011). They also prefer the use of neutral and not controversial topics. Gray (2002) stated that new global ELT textbooks are following similar guidelines to cultural content, which are based on two elements: appropriacy and inclusivity. Inclusivity means a fair and balanced representation of people from different races, religions, genders, ethnicities, and classes, while appropriacy refers to avoiding some topics that might be offensive to the book users. However, more recent research has proven that ELT global textbooks still suffer from bias and stereotype (e.g. Amerian and Tajabadi, 2020; Ansary and Babaii, 2003; Dar & Masroor, 2019; Melliti, 2013; Shah & Pathan, 2016). Reimann (2009) also declares that designing books which engage students with relevant and unbiased information on a cultural level is difficult and time consuming and requires a great deal of research and piloting.

## 2.1. Criteria for cultural content in textbook evaluation

Several scholars propose varied criteria for evaluating cultural content in a textbook. Hatoss (2004) focuses on the evaluation of three dimensions: text and visual input, methods used to teach the cultural content, and aims set by the author/s or publisher/s of a particular textbook for developing learners' cultural competence. Brooks (1986, as cited in Tüm and Sarkmaz, 2012) asserts that the cultural items form greeting, personal perceptions, attitudes on different issues, the use of phrases in different situations, commercial facts and domestic animals, etc.

Byram (1993) is considered to be one of the most prominent scholars, emphasizing the importance of introducing the intercultural element to foreign language learners (Sándorová, 2014). Byram (1993, as cited in Juan, 2010) proposes eight areas focusing on cultural content to be included in textbooks. The list consists of eight categories and each category is further divided into subcategories. These categories are:

1. Social identity and social group (social class, regional identity, ethnic minorities)
2. Social interaction (differing levels of formality; as outsider and insider)
3. Belief and behavior (moral, religious beliefs; daily routines)
4. Social and political institutions (state institutions, health care, law and order, social security, local government)
5. Socialization and the life cycle (families, schools, employment, rites of passage)
6. National history (historical and contemporary events seen as markers of national identity)
7. National geography (geographical factors seen as being significant by members)
8. Stereotypes and national identity (what is "typical" symbol of national stereotypes).

Taking note of Byram's comprehensive and practical criteria, we decided to analyze the textbook in light of this checklist and cover almost every aspect of the culture.

## 2.2. Culture in the common European framework of reference (CEFR)

The Common European Framework of Reference for Languages: Learning, Teaching, Assessment (2001) provides the design of teaching and learning materials, the assessment of foreign language proficiency, and a transparent and comprehensive basis for the elaboration of language syllabuses and curriculum guidelines. CEFR encourages language learning as a means of preserving linguistic and cultural identity of native culture (C1) and target culture (C2), improving communication and mutual understanding, and combating intolerance and racism. It is founded on the conviction that language learning outcomes are likely to benefit internationally if syllabuses and curricula, textbooks and examinations are shaped by a common understanding. Within this respect, world cultures gain importance as intercultural awareness. Besides, a new trend within this frame is to develop learners' critical thinking skills and intercultural competence at language classes implying other world cultures (C3) (Byram, 2000; Byram et al., 2001; Sercu and Bandura, 2005). The idea of C1, C2 and even multicultural aspects into language teaching and learning lies in the core of CEFR.

## 2.3. Review of studies on the cultural content of EFL textbooks

A number of models and approaches have been put forward for textbook evaluation. Various researchers have used content analysis, and more specifically, item frequency analysis in order to evaluate a textbook.

Regarding the intercultural communicative competence principles and components in an EFL textbook, in a recent study, Amerian and Tajabadi (2020) investigated Oxford's New Headway (4th ed.). To this end, they looked into the teachers' opinion on the cultural content of the textbook and their responses to an intercultural communicative competence checklist. It was concluded that there are biases and inadequacies in terms of the presentation and development of intercultural knowledge, attitude, and awareness. Amerian and Tajabadi (2020) mentioned that the textbook has presented the United Kingdom as the land of opportunities and emphasized that such a biased outlook towards cultural instruction do not equip learners with adequate and accurate knowledge required for successful intercultural communication.

Yet in another attempt to explore the cultural content, Putra et al. (2020) examined the manifestation of cultures and intercultural interactions in three English textbooks published in Indonesia. The cultural representation in the textbook was analyzed based on the approach suggested by Yuen (2011). The study clarified that the examined textbooks do not appropriately demonstrate cultural diversity. It was revealed that cultures of Indonesia and English-speaking countries are presented dominantly in the textbooks. Moreover, intercultural interactions are demonstrated in limited and superficial ways. As McKay (2002) states, while cultural content of teaching materials should not be limited to the culture of English-speaking countries, they should incorporate a variety of cultural elements to help learners develop an interest in language learning and to foster learner motivation.

Some other studies have criticized the biased cultural content in some internationally published ELT textbooks. Shah & Pathan (2016) confirmed the dominant presentation of western culture in two ELT textbooks published by Oxford University Press. The study employed Fairclough's (2003) Critical Discourse Analysis as an analytical framework to investigate the cultural presentation in the textbooks. It was found that the textbooks promote western value and culture and contain lessons on western lifestyles, customs and traditions to westernize the learners and train them to accept western culture as acceptable norm.

In another study, Chao (2011) examined the cultural content of an internationally published ELT textbook called New American Inside Out (elementary level/2008). The method of content analysis was applied and the content was codified. This study revealed that the textbook under study has presented culture with more focus on the introduction of western products, persons and perspectives. Chao (2011) stated that the bias in favor of Target Culture (English-speaking countries) is obvious throughout the textbook and less interest has been put on presentation of local and Asian culture.

Some other studies have revealed the unbalanced cultural presentation in textbooks in favor of target or international target cultures with less focus on source culture. Al-Sofi (2018) used Content Analysis and Cortazzi and Jin's (1999) model to investigate the cultural content of an English textbook, *Well Read 1*, published by Oxford University Press regarding source culture, target culture, and international target culture. The findings revealed that the diverse cultural elements are represented with more focus on the target and international target cultures with less interest on the source culture.

Arslan (2016) conducted similar research in which he applied descriptive content analysis to investigate the cultural content and the usage frequency of elements related to native culture, target culture and international culture in EFL textbooks used for 3<sup>rd</sup> and 4<sup>th</sup> grade students in Turkey. The results of the study indicated textbooks were successful in presenting intercultural knowledge. However, native culture items are fewer than target and intercultural items in the textbooks.

Unbalanced culture presentation is not limited to ELT textbooks. It has also been revealed in the IELTS as an internationally recognized English language proficiency test. Very few studies have been done on the cultural bias of the IELTS exam. As an example, Freimuth (2014) conducted cultural content analysis on the reading component of twenty IELTS exams.

A total of sixty reading passages were examined for cultural capital. The study found that, on average, one reading test contained fourteen cultural references including cultural objects and historical settings. He found that the readings referred to 139 places or regions around the world with only five references pertaining to the Middle East and none to the United Arab Emirates where this study was conducted. In another study conducted in Bangladesh by Khan (2006) as an IELTS examiner, the speaking module of the IELTS test was investigated. The examiners were asked to fill out questionnaires and some examiners were interviewed. The study explored and analyzed the terminology, vocabulary, topics and question patterns of the speaking test for any cultural bias. Findings show that there are a number of culturally inappropriate and unfamiliar topics and also some vocabulary items and phrases that tend to confuse the candidates and affect their performance. The analysis of responses from the examiners reveals that candidates in Bangladesh have difficulty with certain questions and tasks in the IELTS speaking test as they tend toward Western culture and norms of behavior that need background knowledge to be understood.

## 2.4. Content analysis

Researchers define content analysis as a quantification process that classifies the content into numerical data and measures the frequency or the appearance of the analysis unit in each category (Bryman, 2012; Cohen, et al., 2007; Silverman, 2011). Historically, this technique is closely linked with quantitative studies and originally designed to deal with large amounts of raw material (Bryman, 2012). As put by Weber (1990), one of the fundamental features of content analysis is that a vast amount of written data is reduced to smaller groups of information or long texts with loads of words are represented by fewer words or expressions.

As mentioned earlier, in this technique, the focus is on establishing precise categories and counting instances, either for existence or frequency (Bryman, 2012). Instances are systematically identified, coded, and grouped before being analyzed quantitatively (Silverman, 2011). Bryman (2012) believes that it is the systematic nature of this tool that affords its objectivity. By clearly specifying the coding and other procedures, content analysis is replicable in the sense that other researchers could reproduce the study. Checklists are examples of pre-coded questions formed prior to the phase of coding and categorizing data. Checklists are of great importance as they help the objective and systematic transformation of each data into a score (Cohen, et al., 2007). Content analysis can be qualitative, quantitative, or mixed (Holsti, 1969, p. 11). Rose et al. (2014) also stated that content analysis can be carried out quantitatively or qualitatively.

### 2.4.1. The process of content analysis

The procedures needed to be followed in carrying content analysis are as follows:

Step 1. Decide what kind of analysis is going to occur. The researcher must decide whether to code for a single word or set of words or phrases.

Step 2. Decide the number of concepts to code for. Concepts and categories need to be created and defined to limit the analysis.

Step 3. Decide whether to code for existence or frequency of a concept. When coding for existence an instance would only be counted once, no matter how many times it appeared. However, when coding for frequency the number of times an instance appears in a text matters. This would give a clue that a certain concept is more emphasized in a text

Step 4. Decide on how you will distinguish among concepts; whether concepts are to be coded exactly as they appear or if they can be recorded even when they appear in different forms.

Step 5. Decide what to do with "irrelevant" information. The researcher must decide whether irrelevant information should be ignored (as Weber, 1990, suggests), or used to reexamine.

Step 6. Code the texts.

Step 7. Analyze the results. (Colorado State University)

Many scholars (Dörnyei, 2007; Weber, 1990; and Zhang & Wildemuth, 2009) emphasize the importance of creating and applying a coding scheme in conducting content analysis as it influences the outcomes of the whole study to the largest extent.

## 2.5. Quantitative and qualitative approaches

Qualitative content analysis is often referred to as "latent level analysis, because it concerns a second-level, interpretative analysis of the underlying deeper meaning of the data"; while the quantitative content analysis is usually described as "manifest level analysis", providing an objective and descriptive overview of the "surface meaning of the data" (Dörnyei, 2007, p. 246). According to Holsti (1969) qualitative method of content analysis is insightful, whereas quantitative method is mechanical. In other words, qualitative content analysis is mainly inductive, as it draws inferences from the examination of topics and themes

and data; on the contrary, quantitative content analysis is considered to be deductive, aimed at testing hypotheses or finding answers to questions based upon theories or previous empirical research (Sándorová, 2014).

Zhang and Wildemuth (2009) state that “qualitative content analysis goes beyond merely counting words or extracting objective content from texts to examine meanings, themes and patterns that may be manifest or latent in a particular text” (p. 308). Despite these differences, it has been highlighted by numerous scholars that, in research practice, the two approaches are often applied in combination (Dörnyei, 2007; Zhang & Wildemuth, 2009). Actually, they mutually support each other’s advantages and strengthen the validity of the findings, thus both contribute to obtain a broader, holistic knowledge about the examined field of study (Dörnyei, 2007).

### 3. Methodology

#### 3.1. Textbook under evaluation

The main material used in this study was *Complete IELTS*, Bands 4-5 preparation textbook by Brook-Hart & Jakeman (2012) and published by University of Cambridge ESOL examinations. It is widely popular and commonly used as a core teaching material, both in public and private institutes, in Iran and abroad. This textbook is designed for students who wish to take the Academic Module of IELTS. The textbook teaches the skills, and language learners need to reach an intermediate level of English (CEF level B1). The Student Textbook contains ten units, each containing sections on each of the four skills in the IELTS exam.

#### 3.2. Instrument

Byram’s (1993) checklist for the cultural content of the textbook was selected to classify the cultural content of the textbook under certain categories.

#### 3.3. Procedure

##### 3.3.1. Data collection procedure

In order to investigate how culture is presented in the textbook, descriptive content analysis was used. The reason for employing this method is its capacity to turn the collected data of the text under study into manageable categories or codes (Byram & Feng, 2004). Two data collection techniques were chosen, namely, item frequency analysis and checklist. Checklist was used to analyze the cultural content of the textbook. Besides, in order to establish the frequency of culture-related items, item frequency analysis technique was utilized. At the first level, to come up with an overall picture of Center and Periphery Countries’ cultural preferences represented in the textbook, the researcher studied the book from the beginning to the end and listened to all excerpts provided in the form of a DVD accompanying the textbook for the listening sections of the textbook. Different sections of the textbook including listening, reading passages, speaking and writing sections were closely examined to find any potential topics including words representing Center or Periphery Countries. The researcher identified indicators that fit into each category mentioned in the Byram’s (1993) checklist. So, the cultural content of the textbook was classified under certain categories. In the following section content analysis procedure and creating coding scheme are explained.

##### 3.3.2. The procedure of content analysis and coding scheme

As part of the content analysis procedure, it is crucial to decide about a coding scheme. Many scholars (e.g. Dörnyei, 2007; Weber, 1990; and Zhang & Wildemuth, 2009) assert that creating coding scheme is of great importance in conducting content analysis as it influences the outcomes of the whole study. One of the most fundamental and important decisions in developing a coding scheme is the definition of the basic unit of text to be classified. “Word, word sense, sentence, theme, paragraph, and whole text” are six commonly used options (Holsti, 1969, p. 116). In this study, analysis was for single or set of words (= 1 concept).

The following steps were undertaken according to the procedural instructions for the content analysis provided on the Writing@CSU Web Site at Colorado State University.

**Step 1.** Analysis was for single or set of words (=1 concept)

**Step 2.** Two different coding schemes were assigned to certain categories concerning the type of cultural referents. The first group of coding schemes was assigned to the type of cultural items following Byram’s (1993) Cultural Content Checklist including eight components as: social identity and social group, social interaction, belief and behavior, social and political institutions, socialization and the life cycle, national history, national geography, stereotypes and national identity (Byram, 1993, as cited in Juan, 2010).

The second group of coding schemes was assigned to the Center Countries' culture-related items (CC) and Periphery Countries' culture-related items (PC). The words or set of words (=1 concept) representing Center Countries' culture were coded as (CC) and the words representing Periphery Countries' culture were coded as (PC).

**Step 3.** The concepts were coded for frequency not existence. If a reference to a foreign culture was repeated more than once in the same passage, it was counted to the number of times it appeared in order to find out which culture is more emphasized.

**Step 4.** Eight components of Byram's (1993) Cultural Content Checklist were used. If there were any references (either with or without an explanation or detailed description) to the items mentioned in the checklist, they were recorded. Also, a single word and its different forms were recorded.

**Step 5.** "Irrelevant" information was ignored as Weber (1990) suggests.

**Step 6.** Cultural content of the textbook was classified and coded under the eight categories and their subcategories found in Byram's (1993) checklist. At this phase, the researcher identified indicators that fit into each category mentioned in the checklist. Then, any references to Center and Periphery countries were coded. Indirect references were also taken into account. If the name of a country was not specified, any other indirect identity, if provided, was taken as clues to determine the referenced country.

**Step 7.** Coded instances were added up and categorized for quantitative analysis. The number of culture-related referents was counted and the frequency and percentage of Center and Periphery countries' culture presentation were calculated. Finally, the researcher drew inferences from the collected data to investigate what is the dominant culture and also the manner of presentation of culture in the textbook.

#### 4. Results

Ten units of the textbook were examined and their references to Center and Periphery Countries were tabulated. Table 1 presents the general overview concerning the number of sections devoted to four basic skills in the textbook.

**Table 1.** Layout of the textbook

Section	Number
Reading	16
Listening	10
Speaking	10
Writing	10

Each unit is observed to have seven sections, four of which are devoted to four basic skills of Reading, Writing, Speaking, and Listening sections. Other sections including Vocabulary and Spelling, Pronunciation, and Key grammar are equally included in each unit.

Four sections of the textbook including Reading, Writing, Speaking, and Listening were examined thoroughly, and references related to Center and Periphery Countries were classified under the suitable categories included in Byram's (1993) checklist. Finally, the frequency and percentage of Center and Periphery Country's culture presentation were calculated.

The research question concerning the dominant culture presented in the textbook was answered based on the results of the quantification of the data. After that, the researcher drew inferences from the text and the collected data to investigate the manner of the presentation of the Centre and Periphery countries' culture. The study findings are summarized in Table 2 below.

**Table 2.** Frequency of Center and Periphery Countries' Presentation

	CC	Percentage	PC	Percentage	Total
Reading	49	39%	76	60%	125
Listening	0	0%	21	100%	21
Speaking	2	7.69%	24	92.30%	26
Writing	6	100%	0	0%	6
<b>Total</b>	<b>57</b>	<b>32%</b>	<b>121</b>	<b>67.97%</b>	<b>178</b>

Note. CC = center countries; PC = periphery countries.

As seen in table 2, 121 out of 178 coded items in the textbook were references to Periphery Countries, which constitutes 67.97% of the data, whereas only 57 items were references to Center Countries, which occupies only 32% of the study data.

As the data makes it clear, 76 out of 125 coded items in the reading passages were evaluated as representation of Periphery Countries, which constitute 60% of the data, this number is almost two times bigger than Center Countries' presentation in the reading passages (39%), and 49 items were direct references to Center Countries, which constitute 39% of the data in reading passages.

In the listening sections, 21 items were categorized in Byram's (1993) checklist as culture representing items. All the collected data in the listening sections were references to Periphery Countries with no item representing Center Countries. Topics presented in the listening sections were mostly general and culture-free. They contained cultural information that could not be restricted only to one nation or country. Actually, they were general enough to be used by people in any country or culture. These sentences had no indicator of their referents. It needs to be mentioned that throughout this study cultural items were coded based on eight categories found in Byram's (1993) checklist. So, when it is said that some items were culture-free it means that they were not suitable to be categorized according to Byram's (1993) categorization.

The percentage of the presentation of Periphery Countries in speaking sections was 92.30%. On the other hand, Center Countries' presentation constitutes only 7.69% of the collected data in this section. In fact, only two items, which constitute 8% of the coded data in speaking sections, were evaluated as references to Center Countries.

Writing sections were evaluated based on the topics presented for the writing tasks. As seen in table 2, 6 items were evaluated as direct references to Center Countries with no reference to Periphery Countries.

The most frequently used analysis categories in Byram's checklist are clear in Table 3 and Table 4.

**Table 3.** Frequency of Center and Periphery Countries' Representation in Eight Categories Found in Byram's Checklist

	CC	Percentage	PC	Percentage	Total	Percentage
Social identity and social groups	0	0%	6	4.95%	6	3.37%
Social interaction	0	0%	0	0%	0	0%
Belief and behavior	2	3.50%	1	1%	3	1.68%
Social and political institutions	19	33.33%	3	2.47%	22	12.35%
Socialization and life cycle	0	0%	3	2.47%	3	1.68%
National history	3	5.26%	1	1%	4	2.24%
National geography	14	24.56%	85	70.24%	99	55.61%
Stereotypes and national identity	19	33.33%	22	18.18%	41	23%
<b>Total</b>	<b>57</b>	<b>32%</b>	<b>121</b>	<b>67.97%</b>	<b>178</b>	<b>-</b>

Note. CC = center countries; PC = periphery countries.

**Table 4.** Frequency of Culture Representation in the Textbook

	Frequency	Percentage
Social identity and social group	6	3.37%
Social interaction	0	0%
Belief and behavior	3	1.68%
Social and political institutions	22	12.35%
Socialization and life cycle	3	1.68%
National History	4	2.24%
National Geography	99	55.61%
Stereotype and national identity	41	23%
<b>Total</b>	<b>178</b>	<b>-</b>

As presented in Table 3 and Table 4, the most frequently used analysis category was category of National Geography. As can be seen, 99 out of 178 coded data in the study were classified under category of National Geography, which constitutes 55.61% of the coded data. In fact, half of the culture presentations in the textbook were references to National Geography. Also, 85 out of 99 coded data under category of National Geography in the textbook were direct references to Periphery



Countries, whereas only 14 items were representation of Center Countries. It means that items classified under category of National Geography in the textbook were mostly representation of Periphery Countries.

Put it in another words, as seen in Table 3, 85 out of 121 coded data representing Periphery Countries in the textbook were direct references to National Geography (70.24%). It can be concluded that Periphery Countries are represented in the textbook mostly through National Geography. On the other hand, only 14 items out of 57 items representing Center Countries were references to National Geography (24.56%). Stereotypes and National Identity goes next with a percentage of 23 in representing culture in the textbook. Moreover, 41 items were classified under this category, out of which 22 items were representation of Periphery Countries and 19 items were references to Center Countries. The presentation of Periphery and Center Countries through Stereotype and National Identity is somehow balanced.

As the data makes it clear in Table 3, each of two categories of ‘Social and Political Institutions’ and ‘Stereotype and National Identity’ constitute 33.33% of the Center Countries’ presentation in the textbook. In other words, references to Center Countries were mostly presented through ‘Social and Political Institutions’ and ‘Stereotype and National Identity’. On the other hand, only 2.47% of the data represents Social and Political Institutions of Periphery Countries.

It needs to be mentioned that, from among twenty-nine subcategories in Byram’s checklist, no data were coded under twelve subcategories in the study. The unused subcategories are: social class, ethnic minority, greeting, moral and religious beliefs, health care, law and order, media, rites of passages, art, music, literature, and food.

## 5. Discussion

Bearing in mind that Complete IELTS textbook is commonly used as a preparation material all around the world, one would expect to witness three cultures -the source, the target and international culture- (McKay, 2002) as equally-represented and to a reasonable extent in the textbook. In the realm of cultural content evaluation of textbooks, many studies found that Center Countries’ culture is represented more prominently than the culture of Periphery countries. The outcome of the current study, to the contrary, suggested that the presentation of culture in the textbook is in favor of the Periphery Countries. These countries made up the largest portion of culture presentation of the textbook under study with 67.97%, while the percentage of presentation of Center Countries is only 32%.

Considering CEFR’s focus on interculturality, the results revealed that the textbook has not met the requirements suggested by CEFR, an important framework in terms of cultural identity and improving intercultural communication. As Complete IELTS textbook is claimed to be prepared according to CEFR, the textbook is expected to have a rich cultural content. It seems that the balance has not been established in terms of including intercultural elements.

As Arslan (2016) mentioned, intercultural communicative competence is regarded an important aim of EFL teaching. Consequently, the investigation of cultural content of textbooks as an important element in the course of teaching and learning has been the focus of many researchers. In this regard, Amerian and Tajabadi (2020) investigated Oxford’s New Headway (4th ed.). The findings proved biases and inadequacies in terms of the presentation and development of intercultural knowledge, attitude, and awareness. They criticized that such a biased outlook towards cultural instruction does not equip learners with adequate and accurate knowledge required for successful intercultural communication.

In a similar vein, Juan (2010) examined the cultural content of EFL textbooks used in higher education in China. She explored the kind of cultural information contained in the college EFL textbooks. It was found that the cultural content input in the textbook has not received attention in designing and organizing the textbooks. She suggested that the comparisons and contrasts between different cultures should be added, more passages should reflect the culture of other English-Speaking communities, and international cultures should be included.

These findings are in accordance with one of the shortcomings found in the analysis of the textbook under investigation in this study. The findings of the present study disclosed the fact that cultural variety has received scant attention in the compilation of *Complete IELTS* textbook. As mentioned earlier, twelve subcategories of the Byram’s checklist were passed over in the process of cultural analysis unused of the textbook. Also, 70.24% of cultural presentation of Periphery Countries in the textbook was references to National Geography and Geographical Items, while Socialization and Life Cycle, Music, Literature, and Social Interaction have received slight or no attention. This is also explicable in light of the link between language and culture. Given the results of the present study and the intricate connection between language and culture, more passages should reflect international cultures to keep learners interested in learning about cultures and developing their critical thinking by comparing and contrasting the cultural values.

The outcome of the present study is explicable in light of the challenges posed against the hegemony of native speakers' norms (Shin et al., 2011). English teaching and learning are encouraged not only to incorporate cultures of English-speaking countries but also involve intercultural perspectives (Baker, 2012a; Baker, 2012b).

It is also in line with the notion of global citizenship. In response to the needs of a global society, individuals should be empowered and placed at center of any educational endeavor. The shift toward student-centered learning is discernible in the recently published EFL textbooks. Meyer et al. (2010) found that measures of student centrism had a positive, robust effect on the level of human rights emphases. In other words, student centrism may contribute to the rise of global citizenship and assist them in getting familiar with the cultural diversity.

## 6. Conclusion

The analysis of the cultural content of the *Complete IELTS* textbook revealed that the textbook's cultural content is not properly inserted. Periphery Countries made up the largest portion of culture presentation in the textbook. In fact, the cultural presentation is more focused on Periphery Countries which constitute 67.97% of the collected data. Center Countries occupied only 32% of the cultural content of the textbook. Some criteria such as National Geography (55.61%), Stereotype and national identity (23%) are presented more significantly, while some other criteria such as socialization and life cycle (1.68%), belief and behavior (1.68%), and social identity and social group (3.37%) are addressed to a lesser degree in the textbook.

Moreover, no item in the book was coded under any of the following criteria in Byram's checklist: greeting, moral and religious beliefs, health care, law and order, media, rites of passages, art, music, literature, and food. The best-case scenario is when all criteria are addressed equally throughout the book by taking cross-cultural learning environments into account in order to familiarize learners with necessary aspects of different cultures and keep them motivated during their learning process. In the present study, however, the majority of the items coded under Periphery Countries were references to the Middle East in the textbook. To sum it up, material designers and textbook writers need to adopt a more critical position towards the cultural content of textbooks. Designing materials presenting cultural diversities and ideologies in different societies can enhance learners' intercultural communicative competence.

## 7. References

- Alkathery, E. R. (2011). *Content Analysis of Culture in ELT Reading Textbooks: Appropriacy and Inclusivity* (Unpublished master's thesis). King Saud University, United Arab Emirates.
- Al-Sofi, B. B. M. (2018). An evaluation of the cultural aspects in the university English textbook, Well Read 1. *Theory and Practice in Language Studies*, 8(2), 184-196. <http://dx.doi.org/10.17507/tpls.0802.02>
- Amerian, M. & Tajabadi, A. (2020): The role of culture in foreign language teaching textbooks: an evaluation of New Headway series from an intercultural perspective. *Intercultural Education*, <https://doi.org/10.1080/14675986.2020.1747291>
- Ansary, H. & Babaii, E. (2003). Universal characteristics of EFL/ESL textbooks: a step towards systematic textbook evaluation. *The Internet TESL Journal*, 8(2). Retrieved from <http://iteslj.org/Articles/Ansary-Textbooks/>
- Arslan, S. (2016). An analysis of two Turkish EFL books in terms of cultural aspects. *Procedia - Social and Behavioral Sciences*. 217-225. <https://doi.org/10.1016/j.sbspro.2016.10.049>
- Baker, W. (2012a). From cultural awareness to intercultural awareness: Culture in ELT. *ELT Journal*, 66(1), 62-70. <https://doi.org/10.1093/elt/ccr017>
- Baker, W. (2012b). Global cultures and identities: Refocusing the aims of ELT in Asia through intercultural awareness. In T. Muller, S. Herder, J. Adamson, & P. S. Brown (Eds.), *Innovating EFL Teaching in Asia* (pp. 23-34). Palgrave Macmillan UK. [https://doi.org/10.1057/9780230347823\\_3](https://doi.org/10.1057/9780230347823_3)
- Bennett, J. M., Bennett, M. J., & Allen, W. (2003). Developing intercultural competence in the language classroom. In D. L. Lange & R. M. Paige (Eds.), *Culture as the core: Perspectives on culture in second language learning* (pp. 237-270). Information Age Pub.
- Brook-Hart, G., & Jakeman, V. (2012). *Complete IELTS Bands 5 – 6.5*. Cambridge University Press.

- Bryman, A. (2012). *Social research methods*. New York: Oxford University Press.
- Byram, M. (1993). Language and culture: The Need for Integration. In M. Byram (Ed.), *Germany: Its Representation in Textbooks for Teaching German in Great Britain* (pp. 3-16). Frankfurt: Diesterweg.
- Byram, M. (1997). *Teaching and assessing intercultural communicative competence*. Clevedon. Multilingual Matters.
- Byram, M. (2000). Assessing intercultural competence in language teaching. *Sprogforum*, 18, 8-13. <https://doi.org/10.7146/spr.v6i18.116742>
- Byram, M., & Feng, A. (2004). Culture and language learning: Teaching, research and scholarship. *Language teaching*, 37(3), 149-168. <https://doi.org/10.1017/S0261444804002289>
- Byram, M., Nichols, A., & Stevens, D. (2001). *Developing intercultural competence in practice*. Clevedon: Multilingual Matters.
- Canagarajah, S. A. (1999). *Resisting linguistic imperialism in English teaching*. Oxford: Oxford University Press.
- Canagarajah, A. S. (2003). *Resisting linguistic imperialism in English teaching*. Oxford: Oxford University Press.
- Chao, T. (2011). The hidden curriculum of cultural content in internationally published ELT textbooks: A closer look at New American Inside Out. *The Journal of ASIA TEFL*, 8(2), 189-210.
- Cohen, L., Manion, L., & Morrison, K. (2007). *Research methods in Education*. New York: Routledge. Retrieved from <http://knowledgeportal.pakteachers.org/sites/knowledgeportal.pakteachers.org/files/resources>
- Cortazzi, M., & Jin, L. (1999). Cultural mirrors: Materials and methods in the EFL classroom. In E. Hinkel (ed.) *Culture in Second Language Teaching and Learning* (pp. 149-176). Cambridge: Cambridge University Press.
- Dar, S., & Masroor, F. (2019). Cultural dissemination: A case study of English textbooks at primary level in Pakistan. *International Journal of Innovation in Teaching and Learning (IJITL)*, 5(2), 33-48. <https://doi.org/10.35993/ijitl.v5i2.699>
- Dornyei, Z. (2007). *Research Methods in Applied Linguistics: quantitative, qualitative and mixed methodologies*. Oxford: Oxford University Press.
- Fairclough, N. (2003). *Analysing discourse: Textual analysis for social research*. London and New York: Routledge Taylor and Francis Group.
- Freimuth, H. (2014a). *Cultural bias on the IELTS examination: a critical realist investigation*. (Unpublished doctoral dissertation). Rhodes University, Grahamstown, South Africa.
- Gao, X. (2006). Understanding changes in Chinese students' uses of learning strategies in China and Britain: A socio-cultural re-interpretation. *System*, 34(1), 55-67. <https://doi.org/10.1016/j.system.2005.04.003>
- Gray, J. (2002). The global course book in ELT. In D. Block & D. Cameron (Ed.), *Globalization and language teaching* (pp. 151-67). London: Routledge. <https://doi.org/10.1002/9781444324068>
- Han, Y., & Bae, Y. (2005). An analysis of the cultural content of high school and college English textbooks. *English Teaching*, 60(4), 47-70.
- Hatoss, A. (2004). A model for evaluating textbooks. [Electronic version]. *Babel* 39(2), 25-32. Retrieved from [http://eprints.usq.edu.au/1137/3/Hatoss\\_Babel\\_2004\\_PV.pdf](http://eprints.usq.edu.au/1137/3/Hatoss_Babel_2004_PV.pdf)
- Hill, M. (2016), The new era of the university press: The critical role of established presses. *Learned Publishing*, 29: 316-319. <https://doi.org/10.1002/leap.1059>
- Holsti, O. R. (1969). *Content analysis for the social sciences and humanities*. MA: Addison-Wesley Publishing Company

- Jamalvandi, B. (2013). Realization of culture in English textbooks in Chinese high school level. *Journal of Pan-Pacific Association of Applied Linguistics*, 16(2), 89-101.
- Juan, W. (2010). A Content analysis of the cultural content in the EFL textbooks. *Canadian Social Science*, 6(5), 137-44. <http://dx.doi.org/10.3968/j.css.1923669720100605.016>
- Khan, R. (2006). The IELTS speaking test: Analyzing cultural bias. *Malaysian Journal of ELT Research*, 2, 60-79. <http://dx.doi.org/10.1177/026553228500200204>
- Kramch, C. (1993). *Context and culture in language teaching*. Oxford: University Press.
- Kramsch, C. (1995). The cultural component of language teaching. *Language, Culture, and Curriculum*, 8, 83-92. <https://doi.org/10.1080/07908319509525192>
- Liddicoat, A., & Scarino, A. (2013). *Intercultural language teaching and learning*. Hoboken: Wiley-Blackwell.
- McKay, S. L. (2002). *Teaching English as an international language*. Oxford: Oxford University Press.
- Melliti, M. (2013). Global content in global course books: the way Issues of inappropriacy, inclusivity, and connectedness are treated in headway intermediate. *SAGE Open*, 1-12. <https://doi.org/10.1177/2158244013507265>.
- Meyer, J. W., Bromley, P., & Ramirez, F. O. (2010). Human rights in social science textbooks: Cross-national analyses, 1970–2008. *Sociology of Education*, 83(2), 111-134. <https://doi.org/10.1177/0038040710367936>
- Mitchell, R., & Myles, F. (2004). *Second language learning theories* (2nd ed.). London: Arnold.
- Phillipson, R. (1992). *Linguistic imperialism* [DX Reader version]. Retrieved from <https://books.google.com/books?id=4jVeGWtzQ1oC&printsec>
- Putra, T. K., Rochsantiningsih, D., & Supriyadi, S. (2020). Cultural representation and intercultural interaction in textbooks of English as an international language. *Journal on English as a foreign language*, 10(1), 163-184. <https://doi.org/10.23971/jefl.v10i1.1766>
- Reimann, A. (2009). A critical analysis of cultural content in EFL materials. *Journal of the Faculty of International Studies*, 28, 85–101.
- Rose, S., Spinks, N., & Canhoto, A. I. (2014). *Management research: Applying the principles*. Routledge.
- Sandorova, Z. (2014). Content analysis as a research method in investigating the cultural components in foreign language textbooks. *Journal of Language and Cultural Education*, 2(1), 95-128.
- Sercu, L., & Bandura, E. (2005). *Foreign language teachers and intercultural competence – an international investigation. Multilingual matters*. Retrieved from: [www.multilingual-matters.com](http://www.multilingual-matters.com) [01.04.2014]
- Shah, S. W. A., & Pathan, H. (2016). Representation of western culture in O’level English language textbooks. *ELF Annual Research Journal*, 18, 23-42.
- Shin, J., Eslami, Z. R., & Chen, W.-C. (2011). Presentation of local and international culture in current international English-language teaching textbooks. *Language, Culture and Curriculum*, 24(3), 253–268. <https://doi.org/10.1080/07908318.2011.614694>
- Silverman, D. (2011). *Interpreting qualitative data*. London: Sage Publications.
- Taki, S. (2008). International and local curricula: The question of ideology. *Language Teaching Research*, 12(1), 127–142. <https://doi.org/10.1177/1362168807084503>
- Tüm, G., & Sarkmaz, Ö. (2012). Yabancı dil türkçe ders kitaplarında kültürel öğelerin yeri, *Hacettepe University, Faculty of Education*, 43, pp. 448-459.

Weber, R. P. (1990). *Basic content analysis*. California: Sage publications.

Yuen, K.-M. (2011). The representation of foreign cultures in English textbooks. *ELT Journal*, 65(4), 458–466. <https://doi.org/10.1093/elt/ccq089>



Zacharias, N. T. (2005). Teachers' beliefs about internationally-published materials: A survey of tertiary English teachers in Indonesia. *RELC Journal*, 36(1), 23-37. <https://doi.org/10.1177/0033688205053480>

Zhang, Y., & Wildemuth, B. M. (2009). Qualitative analysis of content. Retrieved from [https://www.ischool.utexas.edu/~yanz/Content\\_analysis.pdf](https://www.ischool.utexas.edu/~yanz/Content_analysis.pdf)



<https://doi.org/10.22077/ali.2023.7216.1024>

# The Impact of Film Watching on Preschoolers' Language Development: A Comparative Study

Mahdieh Noori<sup>1\*</sup>   
 Khadijeh Lotfi<sup>2</sup> 

\*<sup>1,2</sup> Ph.D. in Teaching English as a Foreign Language (TEFL), Department of English, Faculty of Literature, Alzahra University, Tehran, Iran.

## ARTICLE HISTORY

Received: 26 April 2023  
 Revised: 13 July 2023  
 Accepted: 16 August 2023  
 Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [m.noori@alzahra.ac.ir](mailto:m.noori@alzahra.ac.ir)

## ABSTRACT

This study investigates the impact of film watching versus traditional narrative listening on the language development of Iranian preschoolers across three age groups (2.5-4 years, 4-5.5 years, and 5.5-7 years). The primary objective is to assess and compare language development in preschoolers exposed to narratives through both traditional listening and animated film watching. The study investigates differences in language skills, such as topic maintenance and descriptive abilities, between these two exposure methods. Conducted in Iran, the study encompasses various settings where children were exposed to both narrative listening and animated film watching, followed by story recounting over five sessions. The collected data were meticulously coded to identify disparities in topic maintenance and the length of story descriptions. Key findings reveal that traditional narrative listening resulted in more consistent topic maintenance and longer descriptions compared to film watching. Older children consistently crafted similar stories, demonstrating a deeper understanding of literacy concepts. Additionally, children with more educated mothers exhibited extended and fluent language skills, while others excelled in the context of film watching. In summary, this research underscores the significance of considering different media in preschoolers' language development, highlighting their distinct impacts on language skills. These findings carry significant implications for literacy practices, parenting, childcare, materials development, and policymaking. Ultimately, the study provides valuable insights for teachers, parents, and policymakers, informing the design of effective early childhood language learning strategies and recognizing the role of both traditional and modern media in shaping language development in young children.

**KEYWORDS:** Narrative assessment profile; Narrative culture; Preschoolers; Storytelling; Video description

## 1. Introduction

Storytelling, as a form of authentic performance art, and the oldest form of education (Gottschal, 2013), has been prevalent in all cultures worldwide as to transmit traditions, history, myth, and culture. In simple words, storytelling means expression or retelling of a fiction or nonfiction which includes some sentences and expresses a particular topic (Nava & Pincock, 2011; Padilla, 2013). Storytelling has been a part of everyday communication (Franke et al., 2011), and as a pedagogical form, its efficiency has been proved (Raheim, 2021).

Children from a very early age listen to stories being read at home or nursery schools. Television programs for young children include reading episodes in well-formed linguistic forms which comprise essential story features of setting, episode,

as well as resolution. Hence, young children's competence of these linguistic and narrative features may introduce them conventions of written language long before they learn to read (McCabe & Peterson, 1991).

### 1.1. Storytelling of pre-school children (3-6 years old) in home context

Stories as a linguistic tool reminds one of the past experiences and actions; evaluates and creates new experiences; and helps one to understand the surrounding environment. When stories are (re)told, emotional and social relationships are formed and retained (Fivush & Nelson, 2006; Wang & Fivush, 2005) and hence children are helped out to perceive one's thoughts and actions (Charman & Shmueli-Goetz, 1998).

Moreover, storytelling and retelling can facilitate pre-schoolers' recall of contents and facts (NCTE Committee on Storytelling, 2003), comprehension (Babayiğit et al., 2021), and listening and concentration skills (Isik, 2016; Scott et al., 1995). Storytelling provides an opportunity for nurturing children's creative or artistic expression (Catala et al., 2017), ethical value system development (Rahim & Rahiem, 2012), and more importantly, their later literacy development (Isik, 2016) for which many possible reasons have been enumerated as follows:

1. It can help children begin to reflect on units of language such as phonemes (Incognito & Pinto, 2021), words, and syntax (Adams, 1990).
2. It familiarizes children with the language found in books (Strouse et al., 2018).
3. It provides opportunities to gain experience of how books convey meaning (Strouse et al., 2018).
4. It encourages vocabulary development (Gutiérrez-Clellen, 2002).
5. It models the construction of elaborated informational structures about texts (Gutiérrez-Clellen, 2002).
6. It introduces children to ways of talking about books that they will encounter when they enter elementary school (Dickinson & Smith, 1991).
7. It familiarizes children with print (Mason, 1992).
8. It acts as significant source of new concepts and lessons for young children (Strouse et al., 2018).

### 1.2. Storytelling of pre-school children (3-6 years old) in educational context

Moreover, results of the literature revealed that storytelling may lead to the development of verbal fluency, verbal and nonverbal expressiveness, self-confidence, and teamwork for primary school students (Panc et al., 2015). Storytelling furthermore has been shown to support the development of children's writing, reading, vocabulary, listening skills, and other language abilities (Isik, 2016; Lucarevski, 2016; Miller & Pennycuff, 2008; Ramsey, 2003).

Storytelling and retelling abilities comprise a set of higher-level language, cognitive skills, and abilities namely, the events' sequencing, maintainance of text cohesion by means of explicit linguistic markers, exchange of ideas without support, comprehension of cause-effect relationship, and making a story out of the universal story schemata (Gutiérrez-Clellen, 2002). In other words, narrative skills, as predictors of school success in pre-school or primary schoolchildren, bridge the gap between oral language and literacy through extended, de-contextualized, and cohesive discourse units which children mostly confront in the written texts (Rowe, 2013).

Storytelling requires a macro-structure organization in which the discourse units, unlike conversations, move from topic to topic under more locally negotiated conditions. This macro-structure organization ability is what is demanded by the school discourse situation. In addition, narrative production demands more pre-suppositional judgments in comparison with conversations in which the listener's feedback makes such judgments less crucial. Finally, narration facilitates a meta-linguistic skill in which conscious decisions are made about the most effective ways to relate the tale. In sum, storytelling for young pre-school children in the home context, primary schools, or the educational context can benefit children in multiple ways, which can predict their subsequent or current linguistic and academic achievement (Babayiğit et al., 2021). This achievement level in narrative ability especially seems significant for the preschool children since it can act as an important index of their communicative competence and future academic problems (Smith, 1993)

### 1.3. Children's narrative stages

The development of students' narrative skills does not occur over night but it progresses through a series of stages (Stadler & Ward, 2005). The pre-school years (i.e., 3-6 years of age) seem to be a crucial period in children's development of storytelling abilities. Between the ages of 3 and 4, children become capable story tellers and within the end of their pre-schooling years,

they become able to tell more cohesive and complete stories without a need for the least support or clues (Melzi et al., 2013). With the increase of age, they become progressively more sensitive to the cause-and-effect relationships in the stories.

Below the age of five, children are unable to link pictures into a unified story schemata; they are likely to treat each picture individually (Berman, 1988). By the ages of 5 or 6, they can tell stories with complete plots and a central character (Scott et al., 1995). Older children are capable of incorporating more story grammar components in episodes (Jafari et al., 2012) and to express a hierarchy of story events (Berman, 1988). In their stories, actions are linked by incorporation of cause-and-effect relationships and knowledge of the most related story elements such as goals, characters, and one's attempt to reach a goal (Nielsen et al., 2012). Sophisticated narrative structure, elaboration abilities, and major features of oral conversational language are attained by middle to late elementary school years (Scott et al., 1995). Although young children differ in the extent to which they include different story elements in their story retellings, they seem to be familiar with most of these elements (Trabasso & Broek, 1992).

By the first grade, typically developing students become capable of producing chain narratives, if not true ones. By the end of the fifth grade, they are more able to produce detailed, multiple-episode stories, which contain more complex sentences, well-developed characters, mental state verbs, temporal adverbs, and cause-and-effect relationships between motives and actions (Gillam & Pearson, 2004). Students with low levels of narrative development are more at the risk of prospective academic problems (Paul et al., 1996). Accordingly, more stories are told for the children and more they are asked to retell them, the better their story telling abilities are increased (Charman & Shmueli-Goetz, 1998; Silva et al., 2014).

#### 1.4. Storytelling, retelling, and listening culture in Iran

Narrowing down storytelling, retelling, and listening to the Persian culture, Iran as an ancient country is no exception in having a unique culture, which distinguishes it from other communities. One of the ways in which values and norms of a society are learnt and appropriately internalized is cultural activities, especially storytelling. Iranian culture has been an integration of Iran's geographical location, worldview, wants, wishes, sorrows, feelings, and beliefs and has an Islamic-Persian scent that distinguishes from other cultures (Salehi Amiri & Habibi, 2009).

Story reading or retelling especially in ceremonies has a long tradition in Iran. A part of this dates to Ferdowsi (author of *Shahnameh*) once it appeared as *Shahnamehkhani* (Reading the *Shahnameh*) (Vejdani, 2014). Story retelling or face-to-face story reproduction along with the transfer of managed sensations to the audience has turned into one of the cultural aspects of the East while the West is taking use of them, admitting that story reproduction belongs to the East and particularly, Iran is considered as the epitome of the storytelling coming from the East ('Storytelling has become a major cultural trend in Iran', Alamati, 2023)

One such traditional story retelling in Iran is *Naqqāli* as the oldest form of dramatic performance in Iran, and a great influence on society, from the courts to the villages. The performer – *Naqqāl* – recounts stories in verse or prose accompanied by gestures and movements, and sometimes music and painted scrolls. *Naqqāls* not only act as entertainers but bearers of Persian literature and culture, who need to be acquainted with local cultural expressions, languages and dialects, and traditional music, while improvising with retentive memory and skill to captivate an audience. Until recently, *Naqqāls* were considered to be the most significant bearers of folk-tales, ethnic epics, and Iranian folk music. With the decrease of the popularity of coffeehouses, nomadic tents, houses, and historical venues such as ancient caravanserais where *Naqqāli* was formerly performed, along aging of master performers who were called *morsheds* and a steady drop in the number of skilled *Naqqāls*, as well as popularity of new entertainment, there has been declining interest in *Naqqāli* performance, which has threatened the survival of this dramatic art (Naqqāli, Iranian dramatic story-telling, 2005).

Another especial occasion for storytelling tradition in Iran relates to storytelling, book reading, and *Shahnamehkhani* (Reading the *Shahnameh*) dating back to the winter solstice to educate children and teenage as a pass time especially during the time the social media was not as saturated as today (Shayesteh-Rokh, 2011). In some cities in the past time, those who knew a story would have retold it to others and if a family member can play flute would have accompanied them with singing traditional rhymes (Telkabadi, 1976, cited in Shayesteh-Rokh, 2011). In Damghan (another town in Iran), one of the elderly men would have sung a song and the others would listen to (Shamani damghan, 1999, cited in Shayesteh-Rokh, 2011). As storytelling and listening have a long history in Iran's culture, it would be needed to investigate how today Persian children listen to and retell stories, as it is accomplished in this study.

## 2. Literature Review

As above-mentioned, research in the preschool and school context has revealed beneficial effects of increased opportunities to hear stories for children's literacy growth among many other important features. More specifically, the research conducted in the kindergarten and first-grade class contexts have highlighted this significance. The following researches respectively try to



demonstrate the effect of storytelling for linguistic ability developments of Iranian and non-Iranian typically developing preschoolers. Alongside, the gaps in each area are mentioned as follows.

## 2.1. Storytelling of pre-school children

### 2.1.1. Storytelling of Iranian preschoolers

There are meager studies conducted in the Iranian context on storytelling of/to preschoolers. For instance, Salari's investigation (2002; cited in Jafari, 2010) revealed the beneficial effect of storytelling on the lexicon (retained and comprehensible vocabularies) of 3-6 years-old Iranian children with regards to their age. However, the variable of gender was not found to play a significant role.

Later on, Haji Mohammadi and Rahimian (2010) investigated the ellipsis pattern in the story retelling of 4-6 year-old Persian-speaking children exposed to 60 stories in nursery schools. The results revealed while noun ellipsis was the most-used one, the least one pertained to the verb phrase. Concerning the age variable, the most forms of ellipsis were devoted to the 5-6 years age range. Regarding the gender variable, boys had made more use of ellipsis than their counterpart girls, whilst the latter made more use of clause ellipsis compared to the boys who had more verb phrase ellipsis.

More recently, Mehrabi et al. (2015) assessed the speaking ability of 15 Persian-speaking children (i.e., 4-5, 5-6, and 6-7 years of age) in terms of cohesive and grammatical devices as well as the type and amount of cohesive devices in relation with their age difference. The results revealed that children from various age ranges were able to use cohesive devices; however, with different tendencies. With the increase of age, the quantity of their use of cohesive devices increased but it was not indicative of their ability or tendency to use them. Children in their story retellings had better performance in terms of the number of clauses and different types of cohesive devices they used. This was not observed in their story creation tasks though.

### 2.1.2. Storytelling to preschoolers in other countries

Compared to the Iranian context, larger number of studies are conducted on storytelling of/to preschoolers in other than Iranian context. In this sense, Bayon et al. (2003) investigated the collaborative and integrative virtual storytelling and retelling environments for the 5-7 year-old English, Swedish, and American children. The effects of storytelling on development of creativity, literacy, and developmental attributes were also investigated. The results revealed that using the virtual technological storytelling set-ups, children's motivation is increased as for the development of design partnerships and elaborative storytelling capabilities for the expression of more complex nonlinear stories. Creating such collaborative, supportive, and self-confidence boosting storytelling environment helped children in understanding the initiative events and goals and story improvising in front of an audience.

Isbell et al. (2004) showed how storytelling and story reading influence the language development and story comprehension of young children from 3 to 5 years of age, exposed to the same 24 stories. Both storytelling and story reading were found to positively produce oral language. The results were indicative that young children who heard the stories had better story comprehension in their retelling, while children exposed to story reading improved their language complexity.

Van Kleeck et al. (2011) compared the story retelling of African American and European American kindergartners whose mothers had a high school or less educational level versus a group with mothers of academic degrees. The retold stories were measured based on the Renfrew Bus Story – North American Edition scales of information, sentence length, complexity, and independence. The results revealed systematic effects of maternal education and race on the children's performance. As on all measures, children with educated mothers outperformed those with less-educated mothers. Later on, Catala et al. (2017, June) argued that digital storytelling activities can stimulate creativity in children, as required in the 21st century skills.

More recently, Babayiğit et al. (2021) indicated that language comprehension and narrative skills at 5 years of age directly affect reading comprehension skills and reading gains after explaining general cognitive ability, memory, phonological skills, and mother's education. Also, they demonstrated that listening comprehension influences reading achievement of children. In the same year, Lau and Richards (2021) examined relations between home literacy and preschoolers' development of English as a second language in Hong Kong. Findings demonstrated that the home literacy was positively associated with English vocabulary, letter knowledge, phonological awareness, and word reading skills while child and family characteristics were controlled for.

Additionally, Rahiem (2021) explored the effects of storytelling on early childhood education for Bangladeshi preschoolers. The results indicated that simple digital technology made storytelling more entertaining, engaging, attractive, and communicative. Suggested were the need to increase the ability of teachers to use digital technology; equipping schools with communication technology devices; allocation of governmental funding to modernize school equipment; adjustment of curriculum to meet technological developments, and provision of opportunities for children to learn effective use of technology.

Furthermore, Incognito and Pinto (2021) investigated the relative effect of the child's family context, i.e., parents' occupation and education levels; home literacy; and the school context on literacy skills of 193 preschoolers. The results showed

that parental education level was directly related to preschoolers' performance; lower performance was more seen of children from disadvantaged backgrounds. All participants had significantly increased performance in terms of phonological awareness, notational skills, and textual competence.

In sum, the results of the above-mentioned studies revealed the beneficial effects of the pre-school children's storytelling on the increase of their motivation, expressive elaborative abilities, self-confidence, lexicon, and literacy development. Besides, the educational level of children's mothers was found to be a contributing factor for children's storytelling outperformance.

### 2.1.3. The gaps in the research on the storytelling of the pre-school children

Reviewing the research conducted on the storytelling of preschoolers, there seems to be a positive association between the home literacy environment and monolingual children's language and literacy development (Lau & Richards, 2021). More specifically, it can be concluded that there seems to be a lack of sufficient knowledge about the patterns of book use for storytelling to preschoolers not only in the home but also in the nursery school context.

We also know that there are differences in how the books are read to groups of children. Sometimes, children are engaged in highly interactive discussions that include considerable high-level talk about books whereas in others, the amount of children participation is more circumscribed (Dickinson & Smith, 1991). In addition, the effects of film watching on individual or group storytelling of the pre-school children in home and nursery school contexts have not yet been investigated. Moreover, such a study has not yet been investigated in the Persian context in relation to the Persian narrative reading and listening culture.

The results of the literature revealed that storytelling in front of an audience contributes largely to advancement of language abilities and literacy development of children, and not only their verbal but non-verbal expressiveness as well as other psychological factors such as self-confidence and positive attitude toward teamwork. In sum, the restrictive focus of the literature in this regard means that we have a limited view of the factors contributing to later schooling success, and no sense of the extent to which the home reading experiences are variably or mutually reinforcing.

## 2.2. Significance of the study and research questions

Among the above-mentioned gaps, it should be noted that the research conducted to date on the effect of storytelling of preschoolers has not been plentiful yet. Additionally, most research in this area has mainly focused on children with deficits and not on typically developing children (Babayigit et al., 2021). The literature in general demonstrated the beneficial effects of storytelling on the expressive and literacy skills of children (Charman & Shmueli-Goetz, 1998; Silva et al., 2014), but this has not yet been investigated in terms of these or other possible beneficial or neutral effects from different aspects. Therefore, this gap especially seems a worthwhile consideration for the prospective and current curriculum developments in nursery or primary schools.

More specifically, a quick glance on the research conducted to date on the typically developing preschool children and especially those of the 3-6 years age range reveal a gap on possible effects of storytelling for the children from this life span not only in the Iranian but also in the worldwide context. More importantly, effect of watching short films on the storytelling ability of these children has not been investigated sufficiently either in Iran or worldwide. What complicates the issue furthermore is that nor the effect of book reading and children's subsequent storytelling has been compared with the children's storytelling based on watching short films in the mentioned contexts. Nor has been investigated the effects of film watching on individual or group storytelling of the pre-schoolers in the Persian context in relation to the Persian narrative reading and listening culture. In an attempt to fill these gaps, the researchers of the current study tried to answer the following research questions and test the related hypothesis by investigating the effects of nursery storybook readings and film viewings on Iranian typically developing 3-6 year-old children's storytelling ability:

- 1) What are the features of Iranian preschooler children's storytelling after hearing the story read from the book and the film-watching based on three dimensions of NAP factors (i.e., topic maintenance, event sequencing, length of narration, fluency, event sequencing, story comprehension, and ability in creation of a similar/ related story reveals)?
- 2) Is there any effect of maternal education level on Iranian children's narratives?
- 3) Are there any differences among different Iranian preschooler age ranges (i.e., 2.5-4, 4-5.5, & 5.5-7) in terms of their storytelling after being read and film watching?

Relatedly, there comes the following hypothesis:

- 1) There are differences between the Iranian preschool children's storytelling after hearing the story read from the book and the film-watching,
- 2) Children with higher maternal education can retell longer and more coherent stories,
- 3) Older preschoolers (i.e., 4-5.5, & 5.5-7) can retell better stories,
- 4) Ancient Persian narrative tradition has relation with the Iranian preschoolers' storytelling abilities and tendencies.

### 3. Methodology

#### 3.1. Participants

In the present study, nine typically developing Persian-speaking preschool children (in three age cohorts of 2.5- 4; 4- 5.5; and 5.5- 7) participated in five individual sessions during a two-month period between Feb, 2016 to Apr., 2016. Regarding gender, participants consisted of nearly equal number of males and females (five girls and four boys). All participants had produced their first words before the age of two, as indicated by their mothers. Participants were selected from two non-metropolitan cities in Iran. They were divided into three groups for the purpose of comparison based on age differences: group A (2.5 - 4 years old), group B (4 – 5.5 years old), and group C (5.5-7 years old). The latter seven-year-old group was a preschooler as well. Regarding the Social Economic Status (SES) of mothers, all groups were from middle-class families; two of the mothers had high school diplomas, while the rest had college degrees (A.A., B.A., and PhD candidate levels). Except one mother who was a part-timer language teacher, others were housekeepers then.

#### 3.2. Materials

##### 3.2.1. Shimo rhyming picture books

Maisy Mouse is an animated children's picture book series (Cousins & Nielsen, 1999). These are translated into a Persian rhyming picture book series called *Shimo* by Keshavarz (2015). The series has been popular most probably due to its fabulous rhyme and meter as well as its use of brightly colored pictures.

The stories, which were randomly selected by the researchers, are brought up as follows. The ordering is based on the presentation priority:

1. آب و شکر با لیمو / شربت میسازه شیمو. (*Maisy Makes Lemonade*)
2. شیمو کوچولو می تونه / کار بکنه تو خونه. (*Maisy Cleans up*)
3. آرد و شکر تو سینی / پخته شیمو شیرینی. (*Maisy Bakes a Cake*)
4. شب و کتاب و مهتاب / می گن به شیمو بخواب. (*Maisy's Bedtime*)
5. گاز و کلاچ و دنده / شیمو شده راننده. (*Maisy's Bus*)

##### 3.2.2. Maisy mouse animated series

The videos used for this study included five animated episodes of *Maisy Mouse* series (Cousins & Nielsen, 1999). The series are narrated in English and is not dubbed into Persian. The reason for selection of such animated series related to rhyming picture books was that they were among the bestsellers for young kids over years in Iran ('best sellers for kids', 2023). Each selected episode is related to one of the rhyming picture books and was thus presented with the related story during one session. The five episodes were ordered below according to the session of presentation:

1. Shed
2. Ginger-cat
3. Playhouse
4. Train
5. Playground

#### 3.3. Procedure

The participants were conveniently reached and selected from typically developing Iranian preschool children who were divided into three age groups with equal numbers in each. Of the nine participants, three were three in each age groups: 2.5-4; 4-5.5; and 5.5-7. Every subject had to attend five sessions, each lasting about 30 minutes in the home context. Every session was divided into two phases: a) a read and retell phase and b) a watch and retell phase, each lasting about 15 minutes.

Each session lasted about 30 minutes and included two phases. In the first phase, each researcher recounted one of the five stories selected from a Persian rhyming story book entitled as *Shimo* (Keshavarz, 2015) and recorded each child's retelling of the story. In the second phase, an episode of topic-related animated *Maisy Mouse* series (Cousins & Nielsen, 1999) was presented to the child which was later followed by each child's description of the video (s)he watched.

In the read and retell phase, one of the researchers read a rhyming story to each child. During the story readings, the book illustrations were shown to the child as a visual aid. Then, the child was asked to retell the story back to the researchers, who were both then PhD candidates of TESOL with an efficient qualitative research background. The researchers supported and encouraged child retelling by providing positive feedback, using phrases like "what happened next?" or "and then what?" and open-ended prompts like "and" or "then". After each participant retold the story, he/she was asked a few comprehension questions. The researchers audiotaped the retold story sessions to be transcribed later. Both researchers shared the same language and cultural background as the children, and they were indigenous to the cities where the children lived.

In the watch and retell phase, the child watched one of the five selected episodes of the animated *Maisy Mouse* series. The researcher provided a brief explanation of the video while the children were watching it. The explanations were rough translations of the random English narration to Persian. Then, the child was asked to explain what happened in the video. Encouragements, support, and comprehension questions were provided as well. Child's narration was audiotaped in a way not to distract his/her attention.

The reason behind the selection of maternal education as a variable was to examine the replicability of the findings of Van Kleeck et al., (2011) on systematic effects of maternal education (with academic degrees) on kindergartners' performance story in the Iranian context.

### 3.4. Transcription and coding

Working from the 90 recorded retelling sessions, we as the researchers prepared verbatim transcripts for manual analysis and encoding. While transcribing recorded sessions, subjects' mispronunciations were corrected but their grammatical mistakes were retained for further analysis. The transcripts of various phases were divided into three groups according to each age groups of A, B, and C for comparison.

Coding was done by the NVivo qualitative software (Version 11) based on some pre-determined criteria mainly based on the Narrative Assessment Profile (NAP) (Bliss et al., 1998) (see Table 1). NAP, like most of the available narrative assessment tools is designed to evaluate narrative discourse of those with communicative impairments (Bliss et al., 1998); whilst, its application is not limited to the linguistically-impaired. It was selected because it is flexible and useful in evaluating a variety of discourse features. Among different dimensions of NAP, three factors (i.e., topic maintenance, event sequencing, and fluency) were selected to be of concern in the transcripts. Moreover, the transcripts were analyzed and compared regarding comprehension, length of the narratives, and creation of a similar story.

In order to evaluate topic maintenance, the researchers analyzed the data regarding irrelevant utterances and marked them as violations of topic maintenance. If all utterances were on topic, the transcript was evaluated as having appropriate topic maintenance. Otherwise, it was flagged as inappropriate. Likewise, for assessing event sequencing, occurrence of wrong sequence of events was deemed to be of inappropriate event sequencing. Fluency was assessed by spotting number and duration of pauses, false starts, corrections, and repetitions. Comprehension of the stories was evaluated by both spotting irrelevant utterances in transcripts and analyzing child's answers to the comprehension questions. For narration length, the number of words in an uninterrupted narration was counted; hence, the answers to subsequent prompts or questions were not considered. Finally, the subject's ability in the creation of a similar/ related story was assessed by consideration of the attempt to respond to the researcher's request for telling a similar story. Intra-related coding was checked by another co-author. Any independent creation of even a short similar story was counted as the subject's success in task completion.

**Table 1.** Outline of the investigated variables by each participant and book and film watching stage

		Preschoolers							
Variables		#1	#2	#3	#4	#5	#6	#7	#8
Book #1	Comprehension	✓	✓	✓	✗	✓	✓	✓	✓
	Topic Maintenance	✓	✗	✓	✗	✓	✓	✓	✓
	Event sequencing	✓	✓	✓	✗	✓	✓	✓	✓
	Similar Story	✗	✗	✗	✓	✗	✗	✓	✗
	Length	100	3	19	4	107	80	7	60
Book #2	Comprehension	✓	✗	✓	✗	✓	✓	✓	✗
	Topic Maintenance	✓	✓	✓	✓	✓	✓	✓	✓
	Event sequencing	✓	✓	✓	✗	✓	✓	✓	✓
	Similar Story	✓	✗	✗	✗	✗	✗	✗	✓
	Length	53	63	10	1	110	95	26	146
Book #3	Comprehension	✓	✓	✓	✗	✓	✓	✓	✓
	Topic Maintenance	✓	✓	✓	✓	✓	✓	✓	✓
	Event sequencing	✓	✓	✓	✓	✓	✓	✓	✓
	Similar Story	✗	✓	✗	✗	✗	✗	✗	✓
	Length	47	18	142	4	106	128	41	98
Book #4	Comprehension	✗	✓	✓	✗	✓	✓	✓	✓
	Topic Maintenance	✓	✓	✓	✓	✓	✓	✓	✓
	Event sequencing	✓	✓	✓	✓	✓	✓	✓	✓
	Similar Story	✗	✗	✗	✗	✗	✗	✓	✓
	Length	32	2	88	35	126	111	98	30
Book #5	Comprehension	✗	✓	✓	✗	✓	✗	✓	✓
	Topic Maintenance	✓	✓	✓	✓	✓	✓	✓	✓
	Event sequencing	✓	✓	✓	✓	✓	✓	✓	✓
	Similar Story	✗	✗	✗	✗	✗	✗	✗	✗
	Length	56	20	56	5	101	95	40	97
Film #1	Comprehension	✓	✓	✓	✗	✓	✓	✓	✓
	Topic Maintenance	✓	✓	✓	✓	✓	✓	✗	✓
	Event sequencing	✓	✓	✓	✓	✓	✓	✓	✓
	Similar Story	✗	✗	✗	✗	✗	✗	✗	✗
	Length	6	3	11	17	41	67	23	65
Film #2	Comprehension	✓	✓	✓	✓	✓	✓	✓	✓
	Topic Maintenance	✓	✓	✓	✓	✓	✓	✓	✓
	Event sequencing	✓	✓	✓	✗	✓	✓	✓	✓
	Similar Story	✓	✗	✗	✗	✗	✗	✗	✓
	Length	11	1	66	7	17	120	77	165
Film #3	Comprehension	✓		✓	✗	✓	✓	✓	✓
	Topic Maintenance	✗		✓	✗	✓	✓	✓	✓
	Event sequencing	✗		✗		✓	✓	✓	✓
	Similar Story	✗		✗	✗	✗	✗	✗	✗
	Length	10		10	5	50	111	47	150
Film #4	Comprehension	✓	✗	✓	✗	✓	✓	✓	✓
	Topic Maintenance	✗	✓	✓	✗	✓	✓	✗	✓
	Event sequencing						✓		
	Similar Story	✗	✗	✗	✗	✗	✗	✗	✗
	Length	5	3	18		7	35	15	10
Film #5	Comprehension	✓		✓	✗	✓	✓	✓	✓
	Topic Maintenance	✓		✓	✓	✓	✓	✓	✓
	Event sequencing	✓			✗	✗	✓	✓	✓
	Similar Story	✓		✗	✗	✗	✗	✓	✓
	Length	9		6	7	62	106	37	41

## 4. Results

Findings of assessment of the NAP features including topic maintenance, length of narration, event sequencing, fluency, story comprehension, and ability in creation of a similar/ related story variations will be discussed separately below. The revealed effects of age, form of presentation, and maternal education on story retelling are also brought forth.

### 4.1. Assessed NAP features of storytelling after hearing the story versus film-watching

#### 4.1.1. Topic maintenance

Most subjects maintained topics of the stories in their narration of both films and books. Nevertheless, several transcripts included irrelevant utterances and were marked as inappropriate in terms of topic maintenance during the coding process. Inappropriate topic maintenance was only seen in the transcripts from groups A and B that is in the narration of subjects aged below 5.5 years old. Older subjects from group C maintained the topic consistently during their narration. Subjects generally maintained topics more while retelling book stories than describing films.

#### 4.1.2. Event sequencing

In book retellings all children followed the right sequence of events and younger children succeeded in maintenance of event sequencing like older children. In film descriptions; however, group C (5.5 – 7 years) maintained order of events more than the other two groups.

#### 4.1.3. Fluency

For evaluating fluency, subjects' narratives were coded based on repetitions, false starts, corrections, and pauses, which were all observed and recorded during the transcription process. Surprisingly, more instances of false starts and pauses were observed in the narration of older children, mainly group C. Comparing book retellings and film watchings, narrations of the book stories were more fluent than narrations of film stories for all age groups.

#### 4.1.4. Comprehension

In assessing story comprehension, irrelevant utterances or wrong answers to general comprehension questions were coded as miscomprehension. Almost all subjects from the three age groups fully comprehended the stories and succeeded in answering comprehension questions. However, there were fewer miscomprehensions in book retellings than in film descriptions.

#### 4.1.5. Length

To measure the length of narratives, subjects' uninterrupted narrations were separated. Hence, answers to the researcher's questions or repetitions were not counted. Number of words in an uninterrupted narration of the subject was obtained as the length of narratives.

As expected, group C produced longer narratives than groups A and B in both book retellings and film descriptions. Comparing book retellings and film descriptions for all age groups, an interesting result was that Groups A and B produced longer narratives while retelling book stories whereas group C's narratives of books and films were relatively of equal length.

#### 4.1.6. Creation of similar story

Creation of even a very short story of about two lines was encoded as success in the task. The stories created were either an extension of the book or film stories or a similar story with different characters usually including the subject. Only group C showed consistent ability in the creation of a similar story. Groups A and B failed in this task except in few cases. Success in the creation of a similar story mostly occurred following book readings than after film watching.

## 4.2. Maternal education effects

Children whose mothers had high school diplomas maintained topics more than other children of their age whose parents had university degrees. These children maintained topic throughout their narrations of both films and books consistently while other children with more educated parents had some inconsistencies in topic maintenance in their narratives.

There was no difference in event sequencing between subjects whose mothers were at different educational levels. There also seemed to be an effect of mother's educational level on fluency, since children with less educated parents had more hesitations, false starts and corrections in their narration. Subjects with more educated mothers created longer narratives than subjects in their age groups with less educated mothers while retelling books whereas subjects with less educated mothers produced longer narratives in film descriptions. Regarding mothers' education, there seemed to be a relatively equal level of comprehension between children with differently educated mothers. Regarding the creation of a similar story, there was no difference in this ability between subjects with more or less educated parents of the same age.

### 4.3. Differences in storytelling after being read and film watching of various age groups

The preschooler children aged below 5.5 years old (i.e., groups A and B) had more inappropriate topic maintenance. Older subjects from group C (5.5 – 7 years) maintained the topic consistently during their narration. Whilst all subjects generally maintained topics more during retelling book stories rather than describing films, similarly, group C maintained the order of events more than the other two groups.

Almost all subjects from the three age groups fully comprehended the stories and succeeded in answering comprehension questions. Fewer misconceptions were seen in book retellings than in film descriptions. In this sense, more instances of false starts and pauses were observed in the narration of older children, mainly group C. Comparing book retellings and film watchings, narrations of the book stories were more fluent than narrations of film stories for all age groups.

As expected, group C produced longer narratives than groups A and B in both book retellings and film descriptions. Comparing book retellings and film descriptions for all age groups, Groups A and B produced longer narratives while retelling book stories whereas group C's narratives of books and films were relatively of equal length. Only group C showed consistent ability in the creation of a similar story. Groups A and B failed in this task except in few cases. Success in the creation of a similar story mostly occurred following book readings than after film watching.

## 5. Discussion

In an attempt to answer the previously mentioned research questions, verbatim transcripts were encoded and analyzed and the results in the previous section were obtained regarding differences between book retellings and film descriptions, differences between age groups' narratives and differences between performances of children whose mothers were at different educational levels. In addition, it was examined if narratives can be considered as a cultural medium originating from the ancient Persian tradition.

Comparing retellings of books and film descriptions, more topic maintenance, fluency and comprehension were observed from listening to books compared to film watching. This could be in part due to hearing more words from book readings than from films which included only a few sentences. However, another result was that groups A and B produced longer narratives in retelling books whereas group C created narratives of fairly equal length after book reading and film watching. This points out the advantage of books over films for children aged 2.5 – 5.5 years old. It could probably be implied that books are more effective for linguistic development of children below 5.5 years of age compared to watching cartoon/films. This efficiency of book reading for the Iranian context could also date back to the Persian face-to-face storytelling tradition of storytelling, book reading, and *Shahnamehkhani* (Reading the *Shahnameh*) during the winter solstice to educate children and teenage as a pass time especially during the time the social media was not as saturated as today (Shayesteh-Rokh, 2011). Furthermore, it reminds us of the story retelling or face-to-face story reproduction along with the transfer of managed sensations to the audience as one of the cultural aspects of the East and its epitome, Iran ('Storytelling has become a major cultural trend in Iran', Alamati, 2023).

Generally speaking, storytelling and retelling are fairly advantageous activities for preschoolers and bear several merits including improved comprehension (Isbell et al., 2004), creativity (Agosto, 2016), and memory (Isik, 2016; Vaahtoranta et al., 2019). This study in line with previous studies shows advantages of book reading specifically for preschool children and moreover indicates performance differences of preschoolers and older children in retelling books and films.

Regarding age differences, the elder group C outperformed groups A and B considering topic maintenance, event sequencing and length of narratives; however, unexpectedly, group C's narratives had more pauses, corrections, false starts and repetitions. This result can be explained by the significant difference in the length of narratives between group C and groups A and B. Additionally, only the eldest subjects consistently succeeded in creating a similar story, demonstrating their greater understanding of the literacy process and further implying the inclusion of story reconstruction or extension tasks in preschools' curriculum (Mehrabi et al., 2015).

Unlike the studies by Incognito and Pinto (2021) and van Kleeck et al., (2011) in which children with more educated mothers outperformed other children in all instances, in this study, results were varied. Those with more educated mothers

performed longer and more fluently whereas those with less maternal education performed better in film watching which could imply an effect of social class on children's narrative and literacy skills (Silva et al., 2014).

Despite these findings, which were based on pre-determined criteria, in the transcription and coding process, some noteworthy phenomena were observed which can be the topic of further research. One observation was that most of children tried to maintain a poetic tone and sometimes successfully rhymed their narration in an attempt to sound like the original poetic-like story. This implies that the children paid attention to structure and rhyme as well as content and in attempt to follow the rhyme, they were capable of producing rhyming utterances of their own. This observation was present in narratives of all age groups with different maternal education level. This poetic and rhythmic tendency of the Iranian children in their story retelling further dates back to the poetic and rhythmic face-to-face tradition of storytelling, book reading, and Shahnamehkhani (Reading the Shahnameh) during the winter solstice (Shayesteh-Rokh, 2011) as well as the dramatic story-telling of Naqqāli (Naqqāli, Iranian dramatic story-telling, 2011).

## 6. Conclusion

Children seem to develop storytelling abilities in preschool years. At the age of three to four years old, they gradually become capable storytellers and by the end of schooling, they can tell full stories with the least support or clues (Melzi et al., 2013). The more we tell stories to the child and have him/her retell them, the more child's storytelling abilities develop, which is essential in designing pre-schooling educational programs (van den Broek et al., 2011).

Additionally, in line with the findings of Lau and Richards (2021) regarding the positive link between home literacy-rich environment and children's prospective literacy development, the findings of current study may provide implications for effectiveness of home literacy development of children especially by parents. This could be far more effective when it comes to second/foreign language development. As it was observed here, it was implied that the children paid attention to structure and rhyme as well as content. This observation was present in narratives of all age groups with different maternal educational levels.

As for the implications for educational settings specifically for the Iranian context, story reconstruction or extension tasks should be included in the preschools' curriculum for the children's development of literacy processes. Further, the results provide general guidelines for parents and preschool teachers regarding factors affecting children's literacy development, which should be strengthened during childbearing and child literacy practices. Home literacy practices were also found to be effective for both first and second language development. As for the first language development of Iranian preschoolers, moreover, rhymes were found to be far more effective than film watching. This can provide fruitful implications for the design of literacy and entertaining practices not only in the official preschool contexts but also home environment in Iran. Besides, it may suggest that rhyming book developers and policy makers in Iran invest more in developments of rhyming nurseries in relation to the ancient Iranian storytelling tradition. One limitation of the current study was the convenient selection of preschoolers, due to difficulty of access to them over the long run and in various phases as well as not very large sample accessed. Besides, if further studies can reach preschoolers from more varied cities in Iran, including the metropolitans, it would give more fruitful results.

## References

- Adams, M.J. (1990). *Beginning to read*. MIT Press.
- Agosto, D. E. (2016). Why storytelling matters: unveiling the literacy benefits of storytelling. *Children and Libraries*, 14(2), 21-26. <http://dx.doi.org/10.5860/cal.14n2.21>
- Ahangar, A. & Mojahedi Rezaian (2016). Studying the grounding of verb aspect and voice in Persian speaking children's narratives. *Jostarhaye Zabani (Language Related Research)*, 2(4), 145-167 [In Persian].
- Alamati, h. (2023, Jan 21). Storytelling has become a major cultural trend in Iran', <https://www.javanonline.ir/fa/news/1132221/%D9%82%D8%B5%D9%87%E2%80%8C%DA%AF%D9%88%DB%8C%DB%8C-%D8%AF%D8%B1-%D8%A7%DB%8C%D8%B1%D8%A7%D9%86-%D8%A8%D9%87-%DB%8C%DA%A9-%D8%AC%D8%B1%DB%8C%D8%A7%D9%86-%D8%A8%D8%B2%D8%B1%DA%AF-%D9%81%D8%B1%D9%87%D9%86%DA%AF%DB%8C-%D8%AA%D8%A8%D8%AF%DB%8C%D9%84-%D8%B4%D8%AF%D9%87-%D8%A7%D8%B3%D8%AA>
- Babayiğit, S., Roulstone, S., & Wren, Y. (2021). Linguistic comprehension and narrative skills predict reading ability: A 9-year longitudinal study. *British Journal of Educational Psychology*, 91(1), 148-168. <http://dx.doi.org/10.1111/bjep.12353>



- Bayon, J. R., Wilson, D., Stanton, A., & Boltman, A. (2003). Mixed reality storytelling environments. *Virtual Reality* 7, 54–63. <http://dx.doi.org/10.1007/s10055-003-0109-6>
- Berman, R. (1988). On the ability to relate events in a narrative. *Discourse Processes*, 11, 469–499. <http://dx.doi.org/10.1080/01638538809544714>
- 'Best sellers for kids' (2023, Feb 23). <https://bookland.ir/blog/review/127/%D9%BE%D8%B1%D9%81%D8%B1%D9%88%D8%B4-%D8%AA%D8%B1%DB%8C%D9%86-%DA%A9%D8%AA%D8%A7%D8%A8-%D9%87%D8%A7%DB%8C-%DA%A9%D9%88%D8%AF%DA%A9>
- Bliss, L. S., McCabe, A., & Miranda, A. E. (1998). Narrative Assessment Profile: Discourse analysis for school-age children. *Journal of Communication Disorders*, 31(4), 347-362.
- Catala, A., Theune, M., Gijlers, H., & Heylen, D. (2017, June). Storytelling as a creative activity in the classroom. In *Proceedings of the 2017 ACM SIGCHI Conference on Creativity and Cognition* (pp. 237-242). <http://dx.doi.org/10.1145/3059454.3078857>
- Charman, T., & Shmueli-Goetz, Y. (1998). The relationship between theory of mind, language and narrative discourse: An experimental study. *Cahiers de Psychologie Cognitive (Current Psychology of Cognition)*, 17, 245-71.
- Cousins, L. (Writer) & Nielsen, L. (Director). (1999). *Maisy Mouse* [Animation]. United Kingdom: King Rollo Films.
- Dickinson, D. K. (1990). An environmental approach to informal evaluation in preschools. *Linguistics and Education*, 2, 147-163. [http://dx.doi.org/10.1016/S0898-5898\(05\)80032-4](http://dx.doi.org/10.1016/S0898-5898(05)80032-4)
- Dickinson, D. K., & Smith, M. W. (1991, April). *Styles of book reading in preschool classrooms*. Paper presented at the annual convention of the American Educational Research Association, Chicago, IL.
- Fivush, R., & Nelson, K. (2006). Parent–child reminiscing locates the self in the past. *British Journal of Developmental Psychology*, 24(1), 235-51. <http://dx.doi.org/10.1348/026151005X57747>
- Franke, L., Durbin, C., & Myles, B. S. (2011). Narrative assessment methods in children: a systematic review. AAPC Pub.
- Gillam, R., & Pearson, N. (2004). *Test of narrative language*. Super Duper Publications.
- Gottschal, J. (2013). *The storytelling animal: How stories make us human*. Mariner Books.
- Grugeon, E., & Gardner, P. (2000). *The art of storytelling for teachers and pupils*. David Fulton.
- Gutiérrez-Clellen, V. F. (2002). Narratives in two languages: Assessing performance of bilingual Children. *Linguistics and Education*, 13(2), 175-197. [http://dx.doi.org/10.1016/S0898-5898\(01\)00061-4](http://dx.doi.org/10.1016/S0898-5898(01)00061-4)
- Haji Mohammadi, A., & Rahimian, M. (2010). *The ellipsis pattern in a group of Iranian children's story retelling* (Unpublished master's thesis). Shiraz University, Shiraz, Iran.
- Hickman, R. (2004). Diverse Directions: Visual Culture and Studio Practice. In R. Hickman (Ed) (2004, 2nd ed.). *Art Education 11-18: Meaning Purpose & Direction*. Continuum.
- Hutson-Nechkash, P. (2001). *Narrative toolbox: blueprint for storybuilding*. Thinking Publications.
- Incognito, O., & Pinto, G. (2021). Longitudinal effects of family and school context on the development on emergent literacy skills in preschoolers. *Current Psychology*, 1-11. <http://dx.doi.org/10.1007/s12144-021-02274-6>
- Isbell, R., Sobol, J., Lindauer, L., & Lowrance, A. (2004). The effects of storytelling and story reading on the oral language complexity and story comprehension of young children. *Early Childhood Education Journal*, 32(3), 157-163. <http://dx.doi.org/10.1023/B:ECEJ.0000048967.94189.a3>

- Isik, M. A. (2016). The impact of storytelling on young ages. *European Journal of Language and Literature*, 2(3), 115-118. <http://dx.doi.org/10.26417/ejls.v6i1.p115-118>
- Jafari, S. (2010). The effect of storytelling on the literacy development of children and teenagers. <http://www.kanoonnews.ir/NSite/FullStory/News/?Serv>
- Jafari, S., Agharasouli, Z., Modaresi, Y., & Kamali, M. (2012). Developing a story retelling test for the assessment of language structure in Persian-speaking children. *Bimonthly Audiology-Tehran University of Medical Sciences*, 21(3), 51-61.
- Keshavarz, N. (2015). *Shimo*. Panjereh Publications.
- Lartz, M. N., & Mason, J. (1988). Jamie: One child's journey from oral to written language. *Early Childhood Research Quarterly*, 3, 193-208. [http://dx.doi.org/10.1016/0885-2006\(88\)90022-1](http://dx.doi.org/10.1016/0885-2006(88)90022-1)
- Lau, C., & Richards, B. (2021). Home literacy environment and children's English language and literacy skills in Hong Kong. *Frontiers in Psychology*, 11, 569581. <http://dx.doi.org/10.3389/fpsyg.2020.569581>
- Lucarevski, C. R. (2016). The role of storytelling on language learning: A literature review. *Working Papers of the Linguistics Circle of the University of Victoria*, 26(1), 24.
- Mason, J. M. (1992). Reading stories to children: A proposed connection to reading. In P. Cough, L. Ehri, & R. Treiman (Eds.), *Reading acquisition* (pp. 215-242). Erlbaum. <http://dx.doi.org/10.4324/9781351236904-8>
- McCabe, A. & Peterson, C. (1991). Getting the story: A longitudinal study of parental styles in eliciting narratives and developing narrative skill. In A. McCabe & C. Peterson (Eds.), *Developing narrative structure* (pp. 217-253). Erlbaum.
- Mehrabi Sari, I., Dabir Moghadam, M., & Raghibdoust, Sh. (2015). *The developmental process of cohesive devices in Persian speaking children* (Unpublished doctoral dissertation). Allameh Tabatabaie University, Tehran, Iran.
- Melzi, G., Schick, A., & Bostwick, E. (2013). Latino children's narrative competencies over the preschool years. *Actualidades en psicología*, 27, 1-14. <http://dx.doi.org/10.15517/ap.v27i115.9276>
- Miller, S., & Pennycuff, L. (2008). Power of story: Using storytelling to improve literacy learning. *Journal of Cross-Disciplinary Perspectives in Education*, 1(1), 36-43.
- 'Naqqāli, Iranian dramatic storytelling, intangible cultural heritage, UNESCO' (2011). <https://ich.unesco.org/en/USL/naqqli-iranian-dramatic-story-telling-00535>
- Nava, C., & Pincock, J. (2011). Classroom-based instruction in narration for children learning English: A feasibility study. *Proceedings of the National Conference on Undergraduate Research (NCUR)*, Ithaca Colledge, New York.
- NCTE Committee on Storytelling (2003). *Teaching Storytelling*. [http://www.ncte.org/positions/teaching\\_story\\_telling.shtml](http://www.ncte.org/positions/teaching_story_telling.shtml).
- Nielsen, D. C., Friesen, L. D., & Fink, J. (2012). The effectiveness of a model of language-focused classroom instruction on the vocabulary and narrative development of kindergarten children. *Journal of Education* 192(2). <http://dx.doi.org/10.1177/0022057412192002-309>
- Richards, T., & Richards, L. (2015). NVivo qualitative data analysis software (version 11) [Computer software]. QSR International Pty Ltd.
- Padilla J. (2013). *ACTION! (Active Communication Through Theatrical-Based Intervention of Narratives): Designing a language curriculum* (Doctoral dissertation). California State University, San Marcos.
- Panc, I., Georgescu, A., & Zaharia, M. (2015). Why children should learn to tell stories in primary school? *Procedia - Social and Behavioral Science*, 187, 591 – 595. <http://dx.doi.org/10.1016/j.sbspro.2015.03.110>
- Paul, R., Hernandez, R., Taylor, L., & Johnson, K. (1996). Narrative development in late talkers: Early school age. *Journal of Speech and Hearing Research*, 39, 1295-1303. <http://dx.doi.org/10.1044/jshr.3906.1295>

- Paul, R., Lazlo, C., & McFarland, L. (1992, November). *Emergent literacy skills in late talkers*. Mini-seminar presented at the annual convention of the American Speech-Language-Hearing Association, San Antonio, TX.
- Rahim, H. & Rahiem, M. D. (2012). The use of stories as moral education for young children. *International Journal of Social Science and Humanity*, 2(6), 454-458. <http://dx.doi.org/10.7763/IJSSH.2012.V2.145>
- Rahiem, M. D. (2021). Storytelling in early childhood education: Time to go digital. *International Journal of Child Care and Education Policy*, 15(1), 1-20. <http://dx.doi.org/10.1186/s40723-021-00081-x>
- Ramsey, I. (2003). *Storytelling: Definition and purpose*. <http://falcon.jmu.edu/~ramseyil/storydefinition.htm>.
- Rowe, M. (2013). Decontextualized language input and preschoolers' vocabulary development. *Semin Seech Lang*, 34(4), 260-266. <http://dx.doi.org/10.1055/s-0033-1353444>
- Salehi Amiri, S. R., and Habibi, F., (2009). Evaluating the role of the activities of the intellectual development center for children and adolescents in preserving and strengthening the cultural identity of children and adolescents (based on storytelling), *Journal of Cultural Management*, 3(3), 21-40.
- Scott, L. A., Healey E. C., & Norris, J. (1995). A comparison between children who stutter and their normally fluent peers on a story retelling task. *Journal of Fluency Disorders* 20, 279–292. [http://dx.doi.org/10.1016/0094-730X\(94\)00015-L](http://dx.doi.org/10.1016/0094-730X(94)00015-L)
- Shayesteh-Rokh, E. (2011). Yalda night in the culture of Iranian people. *Culture of Iranian People*. 27(9), 9-26.
- Silva, A., Ferreira, A., & Queiroga, B. (2014). Development of oral narrative and level of mother's education. *CEFAC Journal*, 16(1), 174-186.
- Smith, P. R. (1993). Narrative skills in 4-year-olds with normal, impaired, and late-developing language. *Journal of Speech and Hearing Research*, 36(3), 0022-4685.
- Stadler, M. A., & Ward, G. C. (2005). Supporting the narrative development of young children. *Early Childhood Education Journal*, 33(2), 73-80. <http://dx.doi.org/10.1007/s10643-005-0024-4>
- Strouse, G. A., Nyhout, A., & Ganea, P. A. (2018). The role of book features in young children's transfer of information from picture books to real-world contexts. *Frontiers in psychology*, 9, 50. <http://dx.doi.org/10.3389/fpsyg.2018.00050>
- Torrance, N. G., & Olson, D. R. (1984). Oral language competence and the acquisition of literacy. In A. D. Pellegrini & T. D. Yawkey (Eds.), *The development of oral and written language in social contexts* (pp. 167-181). Ablex.
- Trabasso, T., Stein, N. (1992). Knowledge of goals and plans in the on-line narration of events. *Cognitive Develop*, 7, 133–170. [http://dx.doi.org/10.1016/0885-2014\(92\)90009-G](http://dx.doi.org/10.1016/0885-2014(92)90009-G)
- Vaahoranta, E., Lenhart, J., Suggate, S., & Lenhard, W. (2019). Interactive elaborative storytelling: Engaging children as storytellers to foster vocabulary. *Frontiers in psychology*, 10, 1534. <http://dx.doi.org/10.3389/fpsyg.2019.01534>
- van den Broek, P., Kendeou, P., Lousberg, S., & Visser, G. (2011). Preparing for reading comprehension: Fostering text comprehension skills in preschool and early elementary school children. *International Electronic Journal of Elementary Education*, 4(1), 259-268.
- Van der Pol, C. (2012). Reading picturebooks as literature: Four-to-six-year-old children and the development of literary competence. *Children's Literature in Education*, 43(1), 93-106.
- van Kleeck, A., Lange, A. & Schwarz, A.L. (2011) The effects of race and maternal education level on children's retells of the Renfrew Bus Story — North American Edition. *Journal of Speech, Language, and Hearing Research*, 54, 1546 — 1561. doi: [http://dx.doi.org/10.1044/1092-4388\(2011/10-0079](http://dx.doi.org/10.1044/1092-4388(2011/10-0079)
- Vejdani, F. (2014). Ferdowsi's choice of diction and its role in encouraging anticipation and listening to the public reading of Shahnameh. *Journal of Persian Language and Literature*, 22(77), 275-287. URL: <http://jpll.khu.ac.ir/article-1-2091fa.html>

- Wang, Q., & Fivush, R. (2005). Mother-child conversations of emotionally salient events: Exploring the functions of emotional reminiscing in European-American and Chinese families. *Social Development, 14*(3), 473-95. doi: <http://dx.doi.org/10.1111/j.1467-9507.2005.00312.x>
- Wigren, J. (1994). Narrative completion in the treatment of trauma. *Psychotherapy: Theory, Research, Practice, Training, 31*(3), 415-423. <https://doi.org/10.1037/0033-3204.31.3.415>
- Wright, A. (1995). *Creating Stories with Children*. Oxford University Press.



<https://doi.org/10.22077/ali.2023.7105.1022>

# Translation Assessment of Legal Discursive Structures in the Joint Comprehensive Plan of Action (JCPOA)

Seyed Reza Beh-Afarin<sup>1\*</sup>   
Robabeh Derishemadi<sup>2</sup> 

<sup>1</sup>Assistant Professor in Applied Linguistics, Department of English Translation, Faculty of Foreign Languages, Islamic Azad University, North Tehran Branch, Tehran, Iran

<sup>2</sup>Ph.D. Candidate in Applied Linguistics, Department of English Translation, Faculty of Foreign Languages, Islamic Azad University, North Tehran Branch, Tehran, Iran

## ABSTRACT

This study aimed to examine the translation of legal discourse structures in the Joint Comprehensive Plan of Action (JCPOA) and their impact on the creation, maintenance, and alteration of power relations. The corpus of the study included the written English text of the JCPOA agreement on Iran's nuclear program and its Persian translation. Additionally, the researchers consulted several law dictionaries, including the Law Dictionary published by the Vice Presidency for Legal Affairs (2017), Majd Law Dictionary (2015), and the Legal Terms Dictionary by the Iranian Court of Arbitration for Hague (1989) to compare legal concepts in both languages. Using the Fairclough's model as the analytical framework, the researcher analyzed the original text against its Persian translation. Findings indicated that the Persian translated text (SL) contained some untranslated parts, which was an omission strategy. This omission resulted in translation errors. The discourse practice level error analysis revealed that the Persian translation had an error rate of 11%, which was significant ( $P > 0.05$ ) for Persian translation. On the other hand, the social practice level demonstrated that English terms and words were used in the Persian translation without explanation. The findings were discussed and the implications were made.

**KEYWORDS:** JCPOA (literally in Persian, BARJAM); Translation Assessment; Discursive Structures, Translation Assessment

## ARTICLE HISTORY

Received: 09 April 2023  
Revised: 23 June 2023  
Accepted: 29 August 2023  
Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [r\\_behafarin@iau-tnb.ac.ir](mailto:r_behafarin@iau-tnb.ac.ir)

## 1. Introduction

This research aimed to analyze the translation of legal discourse structures as depicted in the Joint Comprehensive Plan of Action (JCPOA) to assess their significance in establishing, preserving, and modifying power dynamics. Legal discursive structures pertain to the methodologies through which legal language is employed to formulate arguments and convey significance within the legal framework. These structures encompass the utilization of legal terminology, referencing legal precedents and statutes, and employing formalized argumentation techniques like syllogisms and deductive reasoning. Legal discursive structures are structured to guarantee lucidity, exactitude, and uniformity in legal interaction, and they wield a pivotal influence in shaping legal consequences and determinations. This investigation aimed to unpack the complexities of these

discursive structures and explore their impact on power dynamics within the context of the JCPOA. The examination sought to shed light on how legal discourse functions within this framework and its implications for power relations.

The Joint Comprehensive Plan of Action (JCPOA) is a historic and pivotal accord inked on July 14, 2015, in the beautiful city of Vienna, bringing together Iran and the international community, which includes major powers such as China, France, Russia, the United Kingdom, the United States of America, Germany, and the European Union. This landmark pact officially came into force on October 18, 2015, and its provisions commenced on January 16, 2016. At its core, the JCPOA aimed to thwart Iran's potential nuclear weapons program and was crafted as a single agreement with five supplementary appendices. Central to the terms were actions necessitating Iran to dismantle its stockpile of medium-enriched uranium, slash its store of low-enriched uranium by an overwhelming 98%, and curtail its gas centrifuges by approximately two-thirds over a span of 13 years. Moreover, for the subsequent 15 years, Iran consented to limit its uranium enrichment to a mere 3.67% purity, solidifying its commitment to peaceful nuclear energy pursuits. The JCPOA is widely regarded as a considerable diplomatic achievement, embodying years of intricate negotiations and consensus-building.

Iran has committed not to develop any new heavy-water facilities for the following decade. The enrichment activities will be restricted to a single facility utilizing first-generation centrifuges, with other facilities being repurposed to mitigate proliferation risks. The International Atomic Energy Agency (IAEA) will have consistent access to all Iranian nuclear facilities to oversee and validate Iran's adherence to the agreement. In return for demonstrably upholding its obligations, Iran will gain relief from nuclear-related economic sanctions imposed by the US, the European Union (EU), and the United Nations Security Council (UNSC) (Articles 2-6 of the JCPOA's preface, Appendix E). This accord holds significant weight as it ensures the peaceful nature of Iran's nuclear program for the global community. Its translation into Persian is also of critical importance. Historically, Iran holds a substantial background in various international agreements and conventions with nations worldwide, each of which has left a distinct impact on the Iranian government and its people. Notably, the high-stakes oil contracts pre- and post-1979 revolution fall within this realm. This demonstrates Iran's long-standing engagement with international agreements and their implications for the country.

To thoroughly investigate the translation of legal concepts in the JCPOA, the researchers thoroughly examined three law dictionaries: The Vice Presidency for Legal Affairs' law dictionary (2017), Majd Law Dictionary (2015), and the Legal Terms Dictionary by the Iranian Court of Arbitration for Hague (1989). Upon examination, the researchers identified numerous structural and grammatical issues in the Persian translation of the JCPOA's English script. Given the significant implications and importance of agreements like the JCPOA, the translation of legal concepts must be carried out by competent legal experts to guarantee the precision of the translated terms. Furthermore, an accredited translator, sanctioned by the pertinent department of the Ministry of Justice and Foreign Affairs, must translate and officially endorse every document, including the JCPOA agreement. There is no specific evidence to indicate that the JCPOA agreement received official authorization from the General Administration of Documents and Sworn Translators (Technical) Judiciary to confer the same level of validity and support as the original English version in international legal courts. Therefore, the clarity and legitimacy of the Persian translation of the JCPOA may be subject to question.

This research sought to comprehensively analyze the use of Persian equivalents in conveying the precise meaning of English legal concepts within political documents like the JCPOA. Specifically, the study evaluated the effectiveness of Persian translations in preserving the original legal intent as expressed in the English text. Recognizing the high stakes associated with interpreting and translating such documents, the study emphasized the crucial role of legal experts in ensuring faithful rendering. The researchers aimed to communicate the critical importance of accurately translated legal concepts within the JCPOA, acknowledging the significant interpretational impact of legal language in its translation. Employing Fairclough's Model (1993), the study examined the influence of cultural discourse and hegemony on legal concept translation. Recognizing the interconnectedness of Critical Discourse Analysis (CDA) with causality, discursiveness, practice events, and texts (Fairclough, 2013), the study delved into understanding the formation of these practices within a broader social, cultural, and relational framework. Furthermore, it explored how the ambiguity in the discourse-society relationship can be utilized to establish authority and hegemony, highlighting the interplay of power dynamics within the communication of legal concepts.

## 2. Methodology

The study aimed to compare the English text of the JCPOA agreement on Iran's nuclear program with its Persian translation to define and specify legal concepts. The researchers considered this as a subcategory of conceptual research, which aims to clarify concepts. The study focused on two official documents: the written English text of the JCPOA agreement and its Persian translation. The researchers analyzed the content of both documents and consulted several law dictionaries, including the Law Dictionary published by the Vice Presidency for Legal Affairs (2017), Majd Law Dictionary (2015), and the Legal Terms Dictionary by the Iranian Court of Arbitration for Hague (1989).

To compare the Persian equivalents of English legal concepts, the researchers used the principles presented by Fairclough (1993) in his CDA model, which includes causality, discursiveness, practice events, and texts. Fairclough's model emphasizes the social and cultural structures that shape the creation of these practices, events, and texts, and how power

relationships affect the vagueness of discourse and society to secure power and hegemony. The researchers compared the English and Persian versions of the JCPOA agreement to specify legal concepts. However, the accuracy of legal concepts and discursive structures relies on expert legal translators. The study aimed to define these legal concepts and discursive structures and assess the translation.

### 3. Results

The data collected were non-parametric in nature. Therefore, to determine the frequency of discursive structures used in translating legal concepts from English to Persian, a chi-square test was used in the first phase. In the second phase, the researchers analyzed the collected data using three legal dictionaries: (1) Ansari's (2017) Vice Presidency for Legal Affairs law dictionary, (2) Ahsan Nezhad's (2015) MAJD Law Dictionary (14th ed., Vol. 14), and (3) the Legal Terms dictionary by the Iranian Court of Arbitration for Hague (1989). The researchers used these dictionaries to determine the compliance of four lists presented as "proper nouns," "verbs," "adjectives," and "adjective-adverbial phrases" with the original and translated text. In the third phase, the researchers prioritized and identified the most frequently used proper nouns and words in both the original English and Persian translated text to emphasize their importance. According to Fairclough's model of CDA, each language has its own lexical and grammatical aspects, especially in translating legal concepts. It is undeniable that translating from English to Persian requires some shifts at the level of syntax, genres, discourse, and styles. To analyze the translation quality of legal discursive structures in the Joint Comprehensive Plan of Action (JCPOA) and their significance in the production, maintenance, and change of particular relations of power, the researchers collected 100 English adjectives, adverbs, and verbs from the original JCPOA English text.

**Table 1.** List of adjective/adverb verbs & nouns in English and in Persian

A. Adjective/Adverb Verbs, & Nouns in English (COLLOCATIONS)	A. Adjective/Adverb Verbs & Nouns in Persian
1 Historic Joint Comprehensive Plan of Action (JCPOA),	None
2 Be exclusively peaceful	None
3 Positively contribute	None
4 Fundamental shift	None
5 Full implementation	None
6 An exclusively peaceful, indigenous nuclear program	None
7 Mutually determined limitations	None
8 Exclusively peaceful purposes	None
9 Exclusively peaceful nature	None
10 Mutually determined parameters	None
11 Discriminatory regulatory and procedural requirements	None
12 Will be fully observed	به طور کامل رعایت خواهد شد
13 Setting precedents	ایجاد رویه
14 Internationally recognized principles	اصول و رویه های شناخته شده بین المللی
15 Fundamental principles of international law	اصول بنیادی حقوق بین الملل و حقوق و تعهدات
16 Mutually determined civil nuclear cooperation projects	طرح های مربوط به همکاری های هسته ای صلح آمیز که مشترکاً توسط طرفین تعیین می شوند،
17 UN Security Council resolution	قطعنامه شورای امنیت
18 Certain agreed limitations	محدودیت های مورد توافق
19 Enrichment -related activities	فعالیت های مرتبط با غنی سازی
20 Enrichment -related infrastructure	زیر ساخت های غنی سازی مربوطه
21 Safeguarded R&D	تحقیق و توسعه تحت نظارت پادمانی
22 A nuclear, physics and technology Centre	یک مرکز هسته ای، فیزیک و فناوری
23 Appropriate infrastructure modification	اصلاح مقتضی زیر ساخت
24 International qualification standards	استانداردهای بین المللی کیفیت
25 Fabricated fuel assemblies	سوخت تولید شده
26 IAEA technical cooperation	همکاری های فنی آژانس،
27 International qualification standards	استانداردهای کیفی بین المللی
28 IAEA continuous monitoring	نظارت مستمر آژانس
<b>B. ARAK HEAVY WATER REPROCSSING + بازفرآوری آب سنگین</b>	
29 Modernized heavy water research reactor	راکتور تحقیقاتی آب سنگین مدرنیزه
30 Agreed conceptual design	طراحی اولیه مورد توافق

31	Peaceful nuclear research	تحقیقات هسته ای صلح آمیز
32	International technological advancement	پیشرفت فناوری بین المللی
33	Further treatment or disposition	برای پسمانداری یا اقدامات بعدی
34	With the sole exception of separation activities	به جز فعالیت های جداسازی
35	Irradiated enriched uranium targets.	نمونه های تابش دیده اورانیوم غنی شده
<b>C: TRANSPARENCY AND CONFIDENCE BUILDING MEASURES ج. شفافیت و اقدامات اعتمادساز</b>		
36	Comprehensive safeguards agreement	موافقتنامه جامع پادمان
37	With the respective roles	با اختیارات مربوطه
38	Full implementation of activities undertaken	. کاملاً اجرا خواهد کرد
39	Taking necessary action	اتخاذ اقدامات لازم
40	The implementation of the voluntary measures	اجرای اقدامات داوطلبانه
41	To implement transparency measures	اجرای تدابیر شفافیت ساز
42	Long -term IAEA presence	حضور بلند مدت آژانس
43	On -line enrichment measurement and electronic seals;	استفاده از فناوری های مدرن تایید شده و گواهی شده توسط آژانس از جمله دستگاه سنجش میزان غنی سازی به صورت مستقیم و مهر و موم های الکترونیکی
44	The development of a nuclear explosive device	توسعه تجهیزات انفجاری هسته ای
<b>D: SANCTIONS تحریم ها</b>		
45	IAEA-verified implementation of agreed nuclear-related measures by Iran	اجرای اقدامات توافق شده مرتبط با هسته ای توسط ایران، راستی آزمایی شده توسط آژانس،
46	Nuclear -related economic and financial sanctions	تحریم های اقتصادی و مالی مرتبط با هسته ای
47	Including related designations	از جمله فهرست افراد مشخص شده مربوطه،
48	Provision of insurance and reinsurance;	ارائه خدمات بیمه و بیمه اتکالی
49	Specialized financial messaging services	خدمات پیام رسانی مالی
50	Correspondent banking relationships	روابط کارگزاری بانکی جدید
51	Financial assistance and concessional loans	خدمات مالی و وام های ترجیحی
52	Public-guaranteed bonds	اوراق تضمین شده
53	Flagging and classification services	؛(classification) خدمات پرچم و تعیین وضعیت
54	Designation of persons, entities and bodies	لغو فهرست اشخاص حقیقی، حقوقی و نهادها
55	Proliferation -related sanctions	تحریم های مرتبط با عدم اشاعه
56	The agreed nuclear-related measures	اقدامات توافق شده مرتبط با هسته ای
57	Publicly accessible statements	بیانیه هایی را که در دسترس عموم باشد،
58	Consistent with their respective laws	، منطبق با قوانین خود
59	Ongoing investigations	تحقیقات در دست رسیدگی
60	International participants	شرکای بین المللی
61	Agreed joint advanced R&D	تحقیق و توسعه پیشرفته مشترک
<b>IMPLEMENTATION PLAN</b>		
62	Non -binding opinion	برنامه اجرایی
63	Significant non-performance,	نظریه غیر الزام آوری
64	The old UN Security Council resolutions	«عدم پایبندی اساسی»
65	The dispute resolution process specified in this JCPOA ARAK HEAVY WATER RESEARCH REACTOR	قطعهنامه های سابق شورای امنیت سازمان ملل متحد برای طی فرایند حل و فصل اختلاف پیش بینی شده دربرجام راکتور تحقیقاتی آب سنگین اراک
66	The agreed conceptual design	براساس طراحی مفهومی موافقت شده
67	The existing unfinished reactor	راکتور ناتمام موجود
68	Consistent with their respective national laws	هماهنگ با قوانین ملی خود
69	The safe and timely construction	ساخت و راه اندازی ایمن و به موقع
70	Peaceful nuclear research	تحقیقات صلح آمیز هسته ای
71	The agreed conceptual design	. براساس طراحی مفهومی موافقت شده
72	The relevant Iranian regulatory authority	نظام ایمنی ایران
73	Approved final design.	طراحی نهایی تایید شده
74	Relevant funding contributions.	None
75	Subsequent fuel core reloads	بارگذاری های
76	The timely and safe construction.	ساخت ایمن و به موقع
77	Planned radio-isotope production	برنامه تولید رادیو ایزوتوپ ها
78	Reactor operation program	برنامه بهره برداری از راکتور
79	The Fuel Manufacturing Plant	FMP کارخانه تولید سوخت
80	All excess heavy water	تمامی آب سنگین مازاد بر نیاز
81	The Zero power heavy water reactor	راکتور آب سنگین صفر قدرت
82	Necessary technical modifications	تغییرات لازم



83	International technological advancement	پیشرفت تکنولوژی در سطح بین المللی
84	spent fuel reprocessing	بازفرآوری سوخت مصرف شده
85	Spent fuel reprocessing R&D activities	فعالیت های مربوط به تحقیق و توسعه بازفرآوری سوخت مصرف شده
86	Irradiated enriched uranium	اورانیوم غنی شده تابش داده
87	Peaceful industrial purposes.	مقاصد پزشکی و صنعتی
88	Currently operating units	واحدهای در حال تولید
89	Enrichment -related infrastructure,	زیرساختهای مربوط به غنی سازی
90	Single centrifuge machines	تک ماشینهای سانتریفیوژ
91	Internationally established practices	بهترین رویه های بنا شده بین المللی،
92	The Fordow Fuel Enrichment Plant (FFEP)	تاسیسات غنی سازی فردو
93	Stable isotope production activities	فعالیت های خود را در تولید ایزوتوپ های پایدار
94	In agreed areas of research	در زمینه های توافق شده
95	In furtherance of implementation of the JCPOA	در پیشبرد اجرای برجام
96	Under the comprehensive safeguards agreement	تحت موافقت نامه جامع پادمان
97	Access to such locations	دسترسی به چنین محل هایی
98	Satisfactory arrangements	به طریق رضایت بخش
99	Undeclared nuclear materials	فعالیت های هسته ای اعلام نشده
100	Through containment and surveillance	از طریق محدودسازی و مراقبت

**Table 2.** Descriptive data of translation strategies

Translation strategy	A. Adjective/Adverb Verbs & Nouns in English	B. Arak Heavy Water Reprocessing	C. Building Measures Transparency and Confidence	Total
S1	10	4	21	35
S2	5	3	29	37
Paraphrase	0	0	16	16
Omission	10	0	1	11
Total	25	7	71	100

The study found that out of 100 selected English adjectives, adverbs, and verbs, 35 cases (35%) were translated using the (S1) strategy (translating with the same meaning and appearance), 37 cases (37%) using the (S2) strategy (translating with similar meaning and different appearance), 16 cases (16%) using paraphrasing, and finally, 11 cases (11%) using the omission strategy. The results also showed that the frequency of (S2) and (S1) strategies were the highest among the other translation strategies, and a total of 72% of the selected English adjectives, adverbs, and nouns were translated into Persian using (S1) and (S2) strategies. This indicates that the official translation of the original JCPOA aimed to stay as loyal as possible to the original text, with only 16 cases using paraphrasing and 11 cases using omission strategies, which could be more ambiguous and unclear.

Furthermore, out of 224 selected English verbs from a total of around 7250 words shortlisted, up to 180 verbs were translated into Persian, and 44 verbs were omitted, resulting in an incidence rate of 80.3%.

**Table 3.** Descriptive analysis of JCPOA words

	Part A	Part B	Part C
Mean	2.4000	1.4286	1.9296
Median	2.0000	1.0000	2.0000
Std. Deviation	1.38444	.53452	.78055
Variance	1.917	.286	.609
Range	3.00	1.00	3.00
Minimum	1.00	1.00	1.00
Maximum	4.00	2.00	4.00

**Table 4.** Translation strategies of part A: adjective/adverb verbs & nouns in English

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	S1	10	13.9	40.0	40.0
	S2	5	6.9	20.0	60.0
	Omission	10	13.9	40.0	100.0
	Total	25	34.7	100.0	
Missing	System	47	65.3		
Total		72	100.0		

**Table 5.** Translation strategies of part B: Arak heavy water reprocessing

*Part B: Arak Heavy Water Reprocessing*

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	S1	4	5.6	57.1	57.1
	S2	3	4.2	42.9	100.0
	Total	7	9.7	100.0	
Missing	System	65	90.3		
Total		72	100.0		

**Table 6.** Translation strategies of part C: building measures of transparency and confidence

		Frequency	Percent	Valid Percent	Cumulative Percent
Valid	S1	23	31.9	32.4	32.4
	S2	31	43.1	43.7	76.1
	Paraphrase	16	22.2	22.5	98.6
	Omission	1	1.4	1.4	100.0
	Total	71	98.6	100.0	
Missing	System	1	1.4		
Total		72	100.0		

Based on the observations in Table 4, Part A, which analyzes Adjective/Adverb Verbs & Nouns in English, the incidence rate of (S1) was 20%, (S2) was 40%, and Omission was 40%. This indicates that the official Persian translation of BARJAM used (S1) and (S2) translation strategies in 60% of cases, while the Omission strategy was used in the remaining 40%. In Table 5, Part B: Arak Heavy Water Reprocessing, the (S1) strategy was used in 57.1% of cases, and the (S2) strategy was used in 42.9% of cases. No (S3) or (S4) translation strategies were used in this section, making the translation more accurate. In Table 6, Part C: Building Measures Transparency and Confidence, the incidence rate of (S1) was 32.4%, (S2) was 43.7%, Paraphrase was 22.5%, and Omission was 1.4%. The frequencies of adjectives, adverbs, verbs, and nouns used in the original English (JCPOA) and translated Persian script (BARJAM) were analyzed. For this section, all 7250 English words of JCPOA were selected for CDA based on the Fairclough model. From the entire selected sample, 224 cases were English verbs and the remaining 100 selected words were adjectives, adverbs, and nouns. Tables 7 and 8 were prepared separately for verbs and other English words, respectively.

**Table 7.** Frequency tabulations, JCPOA, English-Persian verbs

	Percentage of sample verb	Number of verbs	Frequency percentage of verbs	Total words	Total sample words
English	30.8	224	150	30064	7250
Persian	24.8	180	150	33319	7250
Section		A, B, C			
Total	55.6	404	300	63383	14500

Based on Table 7, the sample volume for both the original JCPOA English text and the Persian official translation was 100%. The frequency percentage of verbs in both texts was the same at 150%. However, only 180 English verbs were translated into Persian, while 44 were omitted, resulting in a deviation of 19.7% between English and Persian verbs.

**Table 8.** Frequency tabulations, JCPOA, English-Persian adjectives/ adverbs/ singular and plural nouns

	Percentage of sample words	Number of words	Frequency percentage of words	Total words	Total sample words
English	13.8	100	150	30064	7250
Persian	12.2	89	150	33319	7250
Section		A, B, C			
Total	26	189	300	63383	14500

Based on Table 8, the frequency percentage of 100 selected English words and their Persian translations was 13.8% and 12.2%, respectively, out of a total of 7250 words. The combined frequency percentage for both English and Persian was 150%. An error analysis conducted at the discourse practice level revealed a significant error rate of 11% in Persian translation ( $P > 0.05$ ). In social practice, English terms and words were used in Persian translation, but there was a significant difference in the discourse knowledge produced by Persian versus English legal discursive structures of JCPOA, as per Fairclough's model of CDA. This variation was 6.00% for English verbs (30.8% English versus 24.8% Persian) and 1.6% for English adjectives/adverbs/nouns (13.8% English versus 12.2% Persian).

Based on Foucault's theory of power relations analysis, power refers to actions that bring about change or direct the behavior of others (such as strategy-making). Power, in this perspective, is the overall structure of actions that influence other possible actions. To evaluate the discourse knowledge of the Persian translation text compared to the original English text, the researchers interviewed a political scientist from the Ministry of Foreign Affairs regarding the effectiveness of word usage, verbs, and active and passive voice sentences. The political scientist concluded that the Persian translation text of the original JCPOA is equal to the original English text based on Fairclough's model of CDA, including syntax, discourse, and style.

According to Foucault, there is a close relationship between power and freedom, and power relations are seen as strategic games between free individuals who may influence the activities of others, despite their lack of power. Foucault believes that power is only exerted on free individuals who can make choices, as power aims to affect human choices and shape their activities. From this perspective, it could be argued that the translated text of the JCPOA underestimates the significance of the JCPOA in the production, maintenance, and change of particular power relations in Iran.

#### 4. Discussion

The results of this study support previous research findings, although they may not align with the results of other investigations. Many studies have explored the linguistic dimension of the JCPOA, including those conducted by Aghagolzadeh (2012), Bánhegyi (2014), Samadi et al. (2011) and Sharififar and Azadi (2016). Aghagolzadeh's (2012) study concluded that to maintain up-to-date knowledge in various aspects of legal fields, including political and commercial, it is necessary to be proficient in different law systems, particularly in terminology. Using general corpus-based knowledge to focus on linguistic expressions can help to benefit from the systemic functional meaning-based level, which is SGF. By considering these features, translators can avoid errors or limit them in their translations. Aghagolzadeh analyzed fifteen translated texts from the English-Persian

exam held by the Iranian Judiciary and found that the most frequent errors were omission, mistranslation, and word choice strategies.

Bánhegyi (2014) found that the multitude of translation approaches and methodologies in Political Discourse Analysis (PDA) is hindering scientific progress in the field, as it makes it difficult to compare and contrast them with one another. Furthermore, the cultural, social, and ideological perspectives of the target audience must also be taken into account. Therefore, unification approaches must be considered in relation to social, political, cultural, historical, hermeneutical, political communication, and contextual features. This will enable the interpretation, evaluation, and analysis of these features under a single theory, while also taking into account the scholar's strategies.

Samadi et al., (2011) researched the problems in translating legal terms from Persian into English. The study examined Altay's (2002) six problematic factors as a criterion and compared them with 239 legal terms chosen and extracted from the Civil Code of Iran, translated by Badrian (2001). They concluded that legal translators benefited from using a paraphrasing strategy, loan words as substitutions, and culture-specific items to have the same influence on the target readers. Sharififar and Azadi (2016) conducted research on investigating ideological manipulation in the translation of political texts, with a case study on the Iranian nuclear talks. The study concluded that the ideology of the translator had affected the translation of BARJAM. Accordingly, the most frequent factor in translating the Persian text was distortion, and the least factor was the manipulation of the expansion strategy.

Karimi (2016) conducted research on the challenges in English to Persian translation of contracts and agreements, with a case study on Iranian English translation students. The study concluded that a lack of efficient knowledge in terminology and syntax caused an essential problem for the students' endeavor to translate legal agreements, contracts, texts, and legal concepts in these. Therefore, legal translations must consider the basis of terminology and the function of the source text (ST) to target text (TT), and legal translators should use practice-oriented knowledge.

The findings of this study are consistent with those of Fernando (2011), who concluded that legal translation is a specific professional translation that requires academic knowledge to avoid conceptual duplications and unite legal thematic elements. Additionally, aptitudes, knowledge, behavior, and efficient translation skills are other important factors. Fernando (2011) introduced two legal translation competence models: the PACTE Group Model, which is based on bilingual, extra-linguistic, instrumental, knowledge about translation, strategic sub-competence, and psycho-physiological components to clarify the strategy for solving problematic situations, and the European Master's in Translation (EMT) Group model, which emphasizes aptitudes, knowledge, behavior, and efficient translation skills. Fernando (2011) concluded that not only training but also the relationship between all these factors enables a process-oriented approach to identify, categorize, and solve problems in legal translation in this field of career. In a study conducted by Azizi and Gorjian (2015), the representation of Iran's nuclear program negotiations in two newspapers, Tehran Times and Los Angeles Times, was investigated. Two reviewers calculated the frequency of discourse strategies to determine the inter-rater reliability index through Pearson Correlation Analysis, resulting in an index of ( $r=.725$ ). The percentages of the structures were then calculated, and Chi-square analysis was used to determine whether the differences were significant. The study found that the two newspapers represented their ideas differently on the same event, using macro strategies to show positive or negative attitudes towards the West or Iran. As a result, the newspapers were significantly different at the macro level. The Los Angeles Times emphasized the positive aspects of the West while downplaying the positive aspects of Iran. At the micro level, there was a significant difference between the frequencies of stability, accusing, authority, threatening, and blaming used as micro-discourse features in both newspapers. For example, a portion of the JCPOA international agreement stated that the E3/EU+3 and the Islamic Republic of Iran welcomed the Joint Comprehensive Plan of Action (JCPOA), which ensured that Iran's nuclear program would be exclusively peaceful and marked a fundamental shift in their approach to this issue. They anticipated that the full implementation of the JCPOA would positively contribute to regional and international peace and security. Iran reaffirmed that it would never seek, develop, or acquire any nuclear weapons under any circumstances. Iran envisioned that the JCPOA would allow it to move forward with an exclusively peaceful, indigenous nuclear program, in line with scientific and economic considerations, in accordance with the JCPOA, and with a view to building confidence and encouraging international cooperation.

In this context, the JCPOA describes initial mutually determined limitations that would be followed by a gradual evolution of Iran's peaceful nuclear program, including its enrichment activities, to a commercial program for exclusively peaceful purposes, consistent with international nonproliferation norms. The E3/EU+3 envisioned that the implementation of this agreement would progressively allow them to gain confidence in the exclusively peaceful nature of Iran's program. The JCPOA established mutually determined parameters, consistent with practical needs, and agreed limits on the scope of Iran's nuclear program, including enrichment activities and R&D. The agreement addressed the E3/EU+3's concerns, including through comprehensive measures providing for transparency and verification. The JCPOA resulted in the comprehensive lifting of all UN Security Council sanctions, as well as multilateral and national sanctions related to Iran's nuclear program, including steps on access in areas of trade, technology, finance, and energy.

## 5. Conclusion

In this study, the researchers aimed to explore the impact of translating political and international high-stakes agreements, such as the JCPOA, on various groups of people, including researchers, scholars, specialists, and learners. The researchers wanted to gain a deeper understanding of how the translation process affects these individuals and how it can be improved to enhance their ability and willingness to translate similar documents in the future. To achieve this goal, the researchers compared the English and Persian texts of the JCPOA and analyzed the use of linguistic elements to create impactful political effects and appeal to scholars and specialists. They found that the use of language in political documents is crucial in conveying the intended message and achieving the desired impact. Furthermore, the study highlights the important role that linguists play in assisting politicians in crafting their formal writings and speeches. Linguists have the expertise to analyze the language used in political documents and speeches and can help politicians convey their message more effectively.

This study sheds light on the importance of language in political and international agreements and emphasizes the need for skilled linguists to assist in the translation and crafting of such documents. The study also highlights the various issues that can arise during the translation process, which can affect the understanding of the source text. These issues include additions, omissions, unresolved references, loss of meaning, and inappropriate linguistic variations such as register, style, and dialect. Additionally, there can be inadequate renderings that affect the expression in the target language, such as spelling, grammar, lexical items, text, and style. These issues can impact the transmission of the main and secondary functions of the source text to the target text.

The study also found that the translators' knowledge of special words, phrases, and principles of political and commercial correspondences had a significant impact on the quality of their translations. Therefore, translation teachers and instructors need to improve their students' knowledge of these texts to enhance their translation quality. It is crucial to inform instructors and teachers about the benefits of improving students' level of knowledge of special properties of legal lexicon in their translation work. The researchers suggest using the rendering of JCPOA as an extraordinary sample of an international agreement in the field of translation studies as well as a fundamental assessment and evaluation factor in the translation sector, based on Fairclough's model from 1993. This suggests that the translation of legal texts is an important area of study for translation students and professionals and that improving their knowledge in this area can lead to better translation quality.

As to the research limitations, this study was limited to only one case; therefore, the findings cannot be generalized. Future studies can explore the translation of JCPOA from fresh analytical models, such as Baker's narrative theory. In-depth comparative analyses of different JCPOA translations (official and unofficial) across languages, focusing on potential discrepancies, ideological manipulation, and cultural adaptations can offer fresh insights. Researchers can also analyze the ethical implications of translator choices and strategies within the context of politically sensitive agreements like the JCPOA.

## 6. References


- Altay, A. (2002). Difficulties encountered in the translation of legal texts: The case of Turkey. *Translation Journal*, 6 (2), 3-8.
- Aghagolzadeh, F., & Farazandeh-pour, F. (2012). The Analysis of English-Persian Legal Translations Based on Systemic Functional Grammar Approach (SFG). *Theory and Practice in Language Studies*, 2(1), 126-131. <https://doi.org/0.4304/tpls.2.1.126-131>
- Ahsan Nezhad, M. E. (2015). *MAJD Comprehensive Law Dictionary - bilingual Persian and English law dictionary*. MAJD Publications.
- Ansari, A. (2017). The Vice Presidency for Legal Affairs (2003, 2009, 2012 ed., Vol. 4th edition). Tehran, Iran: The Vice Presidency of Compliance and Publication of Rules and Regulations. Retrieved from [www.dotic.ir](http://www.dotic.ir)
- Azizi, P. & Gorjian, B. (2015, April 4). A Comparative Study on the Representation of Iran's Nuclear Program Negotiation in Newspapers, Tehran Times vs. Los Angeles Times. [www.ijllalw.org](http://www.ijllalw.org), No. 8 (4).
- Bánhegyi, M. (2014). Translation and political discourse. *Acta Universitatis Sapientiae, Philologica*, 6(2), 139-158.
- Badrian, F. (2001). *Translation Civil Code of Iran*. Rahnama Publication.
- Fairclough, N. (1993). *Critical Discourse Analysis*. Routledge.
- Fairclough, N. (2013). Critical discourse analysis, organisational discourse and organisational change. In *Critical Discourse Analysis* (pp. 347-373). Routledge.

- Fernando, P. (2011). Developing Legal Translation Competence: An Interactive Process-Oriented Approach. *International Journal for Legal Communication*, 5, 7-21. Retrieved from <https://archive-ouverte.ch/unige:16166>
- Karimi, M. H. (2016). Challenges in English to Persian Translation of Contracts and Agreements: The Case of Iranian English Translation Students. *Applied Linguistics and Language Research*, 3(6, 2016), 188-198.
- Legal Terms Dictionary by the Iranian Court of Arbitration for Hague (1989). Gange Danesh Publication.
- Samadi, M., Shomoossi, N., & Rasekh, A. E. (2011). Problem in the translation of legal terms from Persian into English. *The International Journal, Language Society and Culture*, 3, 108-114.
- Sharififar, M., & Azadi, M. B. A. Investigating Ideological Manipulation in Translation of Political Texts Case Study: Iranian Nuclear Talks. *International Journal of Educational Investigations*, 3, (2), 61-71.



<https://doi.org/10.22077/ali.2023.7280.1030>

# A Study on Iranian High School Students' Perception of Cheating in Online Assessments

Hossein Navidinia<sup>1</sup>   
 Mohammad Jalil Zarei<sup>2\*</sup> 

<sup>1</sup>Associate Professor in Applied Linguistics, Department of English Language, Faculty of Humanities and Literature, University of Birjand, Birjand, Iran

<sup>2</sup>M.A. in Applied Linguistics, Ministry of Education, Birjand District, Birjand, Iran

## ABSTRACT

The outbreak of the Covid-19 pandemic has forced many educational institutions worldwide to move from in-person to online classes. Although online technologies do facilitate learning and testing, academic dishonesty in online assessments (OAs) remains an issue of concern since technological devices provide students with ample opportunities to cheat in exams. The purpose of this study was to investigate Iranian high school students' perceptions of cheating in OAs. In addition, the study tried to examine the potential differences between male and female students' perceptions of cheating in OAs. The participants were 214 Iranian high school students. They were asked to answer a questionnaire about academic dishonesty in OAs. The results indicated that more than 80% of the participants had experienced cheating in OAs. "Looking at the teaching materials and copying answers" was mentioned as the most frequent method of cheating. The main reason for cheating was "getting a better score", and the main reason for not cheating was "being morally and socially unacceptable". In addition, almost 70% of the participants did not have a negative attitude toward cheating in OAs. Furthermore, there was not any significant difference between male and female students' perceptions of cheating (Sig.>.05).

**KEYWORDS:** Iranian high school students, Cheating, Online teaching, Online assessment.

## ARTICLE HISTORY

Received: 11 June 2023  
 Revised: 17 August 2023  
 Accepted: 01 September 2023  
 Published: 30 September 2023

## CORRESPONDING AUTHOR

E-mail: [zareii.jalil@gmail.com](mailto:zareii.jalil@gmail.com)

## 1. Introduction

Online teaching/learning and thus online assessment have witnessed a fundamental evolution during the past few years due to the Covid-19 pandemic. With the rise of online assessments (OAs), there is concern over the academic integrity of the exams which embraces ethical values, namely honesty, trust, fairness, respect, responsibility, and courage (Fishman, 2014). These values make an educational community committed to learning, meanwhile, they guarantee the quality of learning and the degree granted to students (Holden et al., 2021). These can be guaranteed by safeguarded assessment, meanwhile, it can be very challenging since there is a lack of control over students' behavior in OAs (Noorbehbahani et al., 2022).

With the compulsory transition from in-person assessment to online, the academic integrity of the tests given to students is of paramount importance for the educational institutions due to the reputation they bring forward. Online

technologies do facilitate learning and testing; however, cheating or academic dishonesty in OAs remains an issue of concern since technological devices provide students with more opportunities to take advantage of their technological devices, e.g. wearable devices such as smart glasses and smart watches, hidden cameras, and even scientific cameras (Curran et al., 2011; Lancaster & Clarke, 2017). Similarly, the likelihood of cheating, more specifically getting answers from friends, is higher during OAs compared to in-person exams (Watson & Sottile, 2010) and there are also paid services where they impersonate examinees (Noorbahani et al., 2022). Therefore, it is vital to conduct serious actions to prevent students from cheating as well as digital monitoring in order to make OAs valid (Fluck, 2019).

To this end, identifying the incentives which hinder examination security helps us with having a broader view, and these incentives are not significantly different in online and in-person exams (Turner & Uludag, 2013). According to Salehi & Gholampour (2021) students' primary reasons for cheating are "uselessness of materials" and "not being ready". Examination security is also affected by factors beyond the mode of education and assessment like the moral attitudes of examinees, the strictness of examiners, and the features of online infrastructures used for testing (Chirumamilla et al., 2020).

Given the importance of OA and its reliability, more studies should be conducted to examine the incentives, types, detection, and prevention of cheating, especially among K-12 learners in online classes considering the paucity of studies carried out in this area. Almost all previous studies in this area have been conducted with tertiary-level students (Ahmadi, 2012; Chirumamilla et al., 2020; Dendir & Maxwell, 2020; Iskandar et al., 2021), while K-12 learners have been the focus of investigation in few studies (Middleton, 2020; Zuo et al., 2021). To the best of the researchers' knowledge, no studies could be found regarding the investigation of K-12 students' perception of cheating in OAs. Therefore, the focus of this study is to investigate the perception of K-12 students toward cheating in OAs. The following questions guide this study:

1. What is the frequency of cheating in OAs?
2. What is the students' perception of the reliability of OAs?
3. What are the common methods of cheating in OAs?
4. What are the students' reasons for cheating in OAs?
5. What are the students' reasons for not cheating in OAs?
6. What are the students' attitudes to cheating in OAs?
7. Is there any significant difference between male and female students' perceptions of cheating in OAs?

## 2. Literature Review

### 2.1. Academic dishonesty

As mentioned earlier, academic integrity can be described as a set of moral and ethical values which build trust between course instructors and students. Therefore, academic dishonesty, so-called academic misconduct or academic fraud, has been defined as a violation of the ethical rules of academic centers (Dyer et al., 2020). Although cheating and academic dishonesty can be used interchangeably, the latter is an umbrella term to which cheating belongs.

There have been several interpretations for cheating, yet one of the most comprehensive definitions can be doing any misconduct regarding tests or even assignments so that it leads to producing fallacious results (Cizek, 2012). What has been suggested unanimously in the literature is that cheating in OAs is inevitable (Dendir & Maxwell, 2020). There should be, as a result, novel ways of proctoring, including the use of webcams and identity authentication to guarantee a safe exam (Xiong & Suen, 2018).

Obviously, cheating can be done in various ways, however, some of them can be done exclusively in online environments. There is a consensus among scholars over cheating opportunities running high in online classes compared to on-site proctored classes (Kennedy et al., 2000; Rogers, 2006; Stuber-McEwen et al., 2009). Christe (2003) maintained that students' collaboration and using unauthorized materials during submission are of the main ways of violating academic integrity in online classes.

As there was an urge to transit from traditional courses to online courses due to the Covid-19 pandemic, cheating habits of students leading to violating academic integrity have changed as well because of not being able to handle all the pressure they faced with that amount of online classes and exams that they had not experienced before (Holden et al., 2021). Despite the growth of academic dishonesty, institutional policies and regulations can afford to mitigate the chances of cheating. McCabe et al. (2002) reported a remarkable correlation between the perceived codes of honor, by both institution staff and students, and the decreased violation of academic integrity. In other words, when students are well-informed of the penalty they might face in case of academic dishonesty, the chances of academic dishonesty will be relatively low.



## 2.2. Online assessment

In the literature, it is known as both online assessment and technology-based assessment which is referred to assessing students' performance as well as their learning in an online environment (Iskandar et al., 2021). There is also another definition given by Yoestara et al. (2020) where students' learning can be measured by either a website or an application. Weleschuk et al. (2019) believe that, OAs should have three main features, including the assessment of students' performance, giving feedback, and facilitating students' learning in an online environment. Of studies conducted in this realm, contradictory findings have been found. On the one hand, students preferred OAs to paper-and-pen exams (Howe, 2020; Petrisor et al., 2016). On the other hand, in studies conducted by Amalia (2018) and Khan and Khan (2019), students favored on-site exams due to the technical challenges they faced during online exams.

Jamil et al. (2012) conducted a questionnaire-based study about teachers' preference for online exams and paper exams. Their questionnaire was mainly concerned with affective factors, adaptability, reliability, and practicality. They found out that teachers mostly viewed online exams more positively than paper exams.

Dermo (2009) has also carried out a study at the tertiary level mainly focusing on finding risks in planning online assessments using six factors: 1) affective factors 2) validity 3) practical issues 4) reliability 5) security and 6) learning and teaching. The findings suggest that the most positive aspect of OAs is the contribution they make to students' learning.

There are studies conducted in the Iranian context yielding empirical evidence on cheating. One study has revealed that there are two factors, internal and external, among Iranian students related to the act of cheating (Ahanchiyan et al., 2016). It is also shown in the study of Bahrami et al. (2015) that 45% of students have tried cheating once. Furthermore, Ahmadi (2012) investigated cheating with English Language as a Foreign Language (EFL) students and realized that the most common methods of cheating are "talking to neighboring students" and "copying one's answers" and the most important incentives were either "not being ready for the exam" or "difficulty level of the exams".

Furthermore, Khamesan and Amiri (2011) conducted a questionnaire-based study at the tertiary level, and they came up with the fact that boys cheat more than girls on exams using neighboring students' exam papers. Almost all of these authors called for official regulations to be passed for cheating.

Many studies have sought cheating and plagiarism in online exams during submitting their exam (Bretag et al., 2019; Kocdar et al., 2018; Owunwanne et al., 2010; Zhang et al., 2018) and on the frequency of cheating and attitude toward cheating in exams at tertiary level (Colnerud & Rosander, 2009; Jamil et al., 2012) but few of them made any attempts to investigate the perception of K-12 students toward OAs. Therefore, this study tries to seek the ideas of K-12 students, both male, and female, regarding OAs.

## 3. Methodology

### 3.1. Participants

The participants in this study were 214 Iranian students from public schools in Birjand, Iran. These students belonged to both junior and senior high schools, and their age ranged from 13 to 18 with the mean of 15.75. Out of 214 students, 138 (64.5) were male and 76.1 (35.5) were female. They were chosen based on convenience sampling. The participants were informed of the purpose of the study, and the confidentiality and anonymity of the findings, and their consent was granted before data collection.

### 3.2. The instrument

The instrument used in this study was designed based on the related literature and interviews with experts. Some items were adapted from Salehi and Gholampour's (2021) study on university students' cheating in onsite exams. The questionnaire was anonymous so the students could answer the items freely. It consisted of 5 parts and 41 items asking high school students' perception of OAs. Part one was about participants' demographic information. Part two, including 2 items, was about the frequency of cheating in OAs. The next part, including one item, asked about the participants' perception of the reliability of online assessment. Part four, which had 6 items, was about the Methods of cheating in OAs. The next part, which includes 18 items, was about Reasons for cheating in OAs. Part six, including 6 items, was about Reasons for NOT cheating in OAs. The last part, which had 8 items, was about Students' attitudes to cheating in online assessments. Below each section of the questionnaire, there was a blank space asking students to add to the items listed in the form based on their opinions or write their own opinions about the item(s). The questionnaire was studied by five experts, and 3 experienced teachers for content validity. The Cronbach's alpha for all the items was .71, which can be considered satisfactory.

### 3.3. Data collection

Data collection took place at the end of the school year (Spring 2022). The purpose of the study was explained to the teachers and the school principals whose permission was required. Before data collection, the questionnaire was briefly explained to the colleagues on how to collect the data. In addition, students were informed of the voluntary and anonymous nature of the questionnaire, and their consent was granted.

### 3.4. Data analysis

The obtained data were analyzed descriptively and inferentially. For descriptive statistics, frequencies and percentages were calculated. For inferential statistics, Independent Sample T-test was used.

## 4. Results

### 4.1. Descriptive statistics

#### 4.1.1. Frequency of cheating in OAs

The first part of the questionnaire was about the frequency of cheating in online exams. As indicated in Table one, surprisingly, 189 (88.3 %) of the participants indicated that they have cheated in online exams.

Furthermore, as indicated in Table 1, among 214 participants, 183 (85.5%) of them believed that they cheated more in online exams compared with the onsite exams.

**Table 1.** Frequency of cheating in OAs

Items	never		rarely		sometimes		often		always	
	F	P	F	P	F	P	F	P	F	P
1. I have cheated in my online exams	25	11.7	56	26.2	58	27.1	30	14.0	45	21.0
2. Compared with onsite exams before, I have cheated more in my online exams	31	14.5	39	18.2	38	17.8	44	20.6	62	29.0

#### 4.1.2. Reliability of OAs

In the next part of the questionnaire, the participants were asked about the reliability of online assessments. As indicated in Table 2, 153 (71.5%) participants believed that OA is “often” and “always” less reliable than onsite assessment.

**Table 2.** Reliability of OAs

Item	never		rarely		sometimes		often		always	
	F	P	F	P	F	P	F	P	F	P
1. I believe that online assessment is less reliable than onsite assessment	17	7.9	18	8.4	26	12.1	32	15.0	121	56.5

#### 4.1.3. Methods of cheating in OAs

In the next part of the questionnaire, the participants were asked about the methods of cheating in OAs. As indicated in Table 3, “looking at the teaching materials and finding/copying answers”, “using notes/summaries written on pieces of paper to find the answers”, and “Talking to my classmates through phone, social media, or other communication devices” were the most frequent methods of cheating in OAs.

**Table 3.** Methods of cheating in OAs

Items	never		rarely		sometimes		often		always	
	F	P	F	P	F	P	F	P	F	P
1. Looking at the teaching materials and finding/copying answers	37	17.3	47	22.0	50	23.4	36	16.8	44	20.6
2. Using notes/summaries written on pieces of paper to find the answers	43	20.1	61	28.5	31	14.5	47	22.0	32	15.0
3. Taking the test with my classmates in the same place so we could help each other	132	61.7	17	7.9	23	10.7	14	6.5	28	13.1
4. Talking to my classmates through phone, social media, or other communication devices	78	36.4	45	21.0	21	9.8	25	11.7	45	21.0
5. Sharing the answers with classmates on social media platforms such as WhatsApp, Telegram, etc.	85	39.7	39	18.2	28	13.1	22	10.3	40	18.7
6. Asking others to take the test instead of you	186	86.9	10	4.7	5	2.3	8	3.7	5	2.3

#### 4.1.4. Reasons for cheating in OAs

In the next part of the questionnaire, the students were asked about the reasons for cheating in OAs. As indicated in Table 4, “getting a better score”, “the difficulty of the exam” and “the time limitation in the exam” were mentioned as the most frequent reasons for cheating in OAs.

**Table 4.** Reasons for cheating in OAs

Items	never		rarely		sometimes		often		always	
	F	P	F	P	F	P	F	P	F	P
1. Not being ready for the test	58	27.1	34	15.9	57	26.6	29	13.6	36	16.8
2. Not having enough time for studying	78	36.4	32	15.0	35	16.4	34	15.9	35	16.4
3. Being stressed at the time of the exam	56	26.2	34	15.9	54	25.2	23	10.7	47	22.0
4. The difficulty of the exam	32	15.0	18	8.4	46	21.5	46	21.5	72	33.6
5. The time limitation in the exam	34	15.9	20	9.3	51	23.8	46	21.5	63	29.4
6. Not having enough motivation to study	121	56.5	21	9.8	20	9.3	26	12.1	26	12.1
7. To enjoy it	157	73.4	17	7.9	11	5.1	8	3.7	21	9.8
8. Pressures or persuasion from classmates	157	73.4	24	11.2	14	6.5	11	5.1	8	3.7
9. No severe punishment for cheating	149	69.6	30	14.0	16	7.5	11	5.1	8	3.7
10. The same behavior with the cheaters and noncheaters	78	36.4	32	15.0	26	12.1	16	7.5	62	29.0
11. The bulkiness of the materials	37	17.3	34	15.9	42	19.6	31	14.5	70	32.7
12. The uselessness of the materials	64	29.9	34	15.9	38	17.8	22	10.3	56	26.2
13. The weakness of managing and organizing the exam	73	34.1	55	25.7	28	13.1	25	11.7	33	15.4
14. Lack of having access to teaching materials	109	50.9	50	23.4	24	11.2	13	6.1	18	8.4
15. Not liking teachers	120	56.1	29	13.6	19	8.9	13	6.1	33	15.4
16. Getting a better score	22	10.3	24	11.2	34	15.9	36	16.8	98	45.8
17. Assuming that everybody else will cheat in the exam	48	22.4	25	11.7	29	13.6	32	15.0	80	37.4
18. Being sure that no one will know about it	81	37.9	56	26.2	17	7.9	24	11.2	36	16.8

#### 4.1.5. Reasons for not cheating in OAs

In the next section, the participants’ perception was asked about the reasons for not cheating in OAs. As shown in Table 5, “being morally and socially unacceptable”, and “affecting the rights and scores of other students” were mentioned as the most important reasons for not cheating in OAs.

**Table 5.** Reasons for not cheating in OAs

Items	never		rarely		sometimes		often		always	
	F	P	F	P	F	P	F	P	F	P
1. Because it affects the rights and scores of other students	75	35.0	28	13.1	30	14.0	40	18.7	41	19.2
2. Because it is not religiously acceptable	86	40.2	21	9.8	41	19.2	33	15.4	33	15.4
3. Because it may bring shame and dishonor	69	32.2	36	16.8	41	19.2	39	18.2	29	13.6
4. Because I am afraid of cheating	102	47.7	27	12.6	36	16.8	31	14.5	18	8.4
5. Because it is not possible to cheat in online exams	65	30.4	36	16.8	35	16.4	34	15.9	44	20.6
6. Because it is morally and socially unacceptable	107	50.0	44	20.6	30	14.0	15	7.0	18	8.4

#### 4.1.6. Students attitude toward cheating in OAs

The last part of the questionnaire asked about students' attitudes toward cheating in OAs. As indicated in Table 6, just 30.22 percent of the participants believed that cheating is not right and acceptable in OAs. This percentage was obtained by summing all the percentages for "strongly agree" and "agree" from all 8 items (item 8 had reverse scoring) and dividing the outcome by 8. This means that almost 70% of the participants did not have negative attitudes toward cheating in OAs.

**Table 6.** Students' attitude toward cheating in OAs

Items	Strongly disagree		Disagree		No idea		Agree		Strongly Agree	
	F	P	F	P	F	P	F	P	F	P
1. Cheating is not right, even if the exam is difficult	58	27.1	47	22.0	44	20.6	30	14.0	35	16.4
2. Cheating is not right, even if there is a chance I might fail	98	45.8	28	13.1	35	16.4	23	10.7	30	14.0
3. Cheating is not right, even if it does not affect other students' scores	83	38.8	49	22.9	38	17.8	24	11.2	20	9.3
4. Cheating is wrong, even if the professor has not taught the relevant materials during the semester	70	32.7	27	12.6	44	20.6	35	16.4	38	17.8
5. Cheating is not the right thing to do, even if the professor is not fair in scoring papers	63	10.0	106	16.9	173	27.5	129	20.5	158	25.1
6. Cheating is not right, even if I do not have enough time to study	75	35.0	48	22.4	38	17.8	23	10.7	30	14.0
7. Cheating is not right, even if all students do it	86	40.2	32	15.0	45	21.0	18	8.4	33	15.4
8. Cheating is acceptable considering the low quality of teaching in my online classes	35	16.4	46	21.5	54	25.2	28	13.1	51	23.8

## 4.2. Inferential statistics

### 4.2.1. Gender and attitude toward cheating in OAs

In order to analyze the data inferentially, a T-test was employed to examine the potential differences between male and female participants' attitudes toward cheating in OAs. Table 7 shows the descriptive statistics for male and female answers, and Table 8 shows the results of the T-test. As indicated in Table 8, there is not any significant difference between male and female participants' attitudes to cheating in OAs (Sig>.05).

**Table 7.** Descriptive statistics for male and female participants

		Group Statistics			
		Gender N	Mean	Std. Deviation	Std. Error Mean
Gender attitude	Male	140	20.2786	8.26631	.69863
	Female	74	21.2297	8.09794	.94137

**Table 8.** Independent sample T-test of differences between male and female attitudes

	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference
Gender attitude	Equal variances assumed	-806	.212	.421	-.95116

## 5. Discussion

The first research question of the study endeavored to probe the frequency of cheating in OAs. The results are appalling since about 88% of students admitted they had cheated in online exams. Over 85% believed they cheated in online exams more than onsite ones. Only 25 students out of 214 stated they have never cheated in their OAs. The results corroborate the previous studies' results where cheating was showed to be widespread (Saleh & Meccawy, 2021; Salehi & Gholampour, 2021; Sims, 1993).

There might be some reasons for this finding, including lack of supervision, frequent absence in online classes or impersonation, and parents' pressure on students expecting them good grades. As suggested in Watson and Sottile's (2010) study, the Internet plays an inevitable role in giving students more opportunities to cheat while being low-profile. While taking exams, it aids them to open different websites searching for the correct answers, and thus plagiarizing the content and submitting them in the exams. As for the second reason, the students' tendency toward cheating is mainly due to the ease of impersonation and faking their absence in online classes which result in violating academic integrity as reflected in Diego (2017). Everything gets worse when parents' pressure on students for achieving higher grades is added to the aforementioned motives, it pushes students toward cheating in order not to be ashamed before their parents (Saleh & Meccawy, 2021).

The reliability of online exams is the next concern of this study. Almost 71% of students mentioned OA is "often" and "always" less reliable than onsite exams. At the tertiary level, it is because of a different reasons including the lack of safeguarding measures as a result of poor infrastructures such as poor connectivity, hardware, software, power supply, and lack of online and physical systems (Tuah & Naing, 2021), and these factors are available in K-12 classes as well.

In line with studies already enquiring the cheating methods (Bretag et al., 2019; Colnerud & Rosander, 2009; Kocdar et al., 2018; Zhang et al., 2018), "looking at teaching materials and finding/copying answers", "using notes/ summaries written on pieces of paper to find the answers", and "talking to my classmates through phone, social media, or other communication devices" are of the most common ways of cheating among students.

In terms of reasons for cheating in online exams and the convenience which the Internet offers to students, it is of no surprise that cheating appears to be plausible, especially when it is accompanied by the thirst for getting a better score. The lack of proctorship and the temptation for being better than peers due to the competitive atmosphere in the classroom pave the way for a better score effortlessly (Kennedy et al., 2000; Rogers, 2006; Xiong & Suen, 2018). Similarly, the exams being difficult is the other reason which made students cheat which is most probably due to not being ready for online exams (Ahmadi, 2012). Most students take advantage of the situation where they can do their homework and present it to the teachers without spending much time. Given this, it seems that students are not well-prepared for the exams, hence, they consider their online exams difficult. Similarly, time limitation gives weight to their cheating as it is recommended in the literature for teachers and course instructors to allow students just for one-time attempt for giving response to the questions (Wahid & Farooq, 2020).

As for their attitude toward cheating, the findings are shocking. Approximately 30% of students had a negative attitude toward cheating in online exams. Since we asked them explicitly if they had already cheated in online exams and 88% experienced cheating, it can be implied that students had committed the act of cheating after a fashion. The findings of previous research questions which indirectly probed their attitude also prove the fact of frequent cheating in online exams.

The nominal demographic variable of gender was also found not to affect cheating behavior among students. In other words, students' desire to get a better score in online exams causes them to cheat irrespective of being male or female. The results of this section contradict those of Jensen et al. (2002) as they found out cheating is more prevalent among male students than female ones. They believed this behavior can be because of the risk-taking characteristics of men which is more than women. However, in another study, Ahmadi (2012) investigated the differences between male and female students' perceptions of cheating in OAs, and his findings showed no significant difference between their perceptions.

## 6. Conclusion

This study endeavored to investigate the perception of the Iranian EFL high school students' perception of cheating in OAs. Due to the findings that prevailed in this study, namely the high frequency of cheating, the poor reliability of online exams, the methods of cheating, the reasons for cheating and not cheating, and the attitude of students to cheating, it is incumbent upon every school and academic institution to publish an easy-to-access ethical guideline explaining dos and don'ts of its online assessments. Cheating is an unforgivable act and there is always the chance of detecting the tinges of cheating in one's submitted response. Students, therefore, should be informed of the penalties that cheating precedes (Ma et al., 2013), and some restrictive conducts such as closing critical ports can be of much help to instructors to safeguard the exams (Rowe, 2004).

Concerning the lack of reliability in OAs, raising teachers' awareness of multiple ways of cheating, obliging students to take their exams in a low-resource setting, and asking them to turn on their webcams are some of the ways which can reinforce the academic integrity of the online exams. Not only does raising awareness help teachers but it suppresses students' tendency toward cheating because of the ethical dilemma it makes so that it will poke their conscience and make them think about the right of others. As Salehi and Gholampour (2021) stated, the weight that social stigma carries is more than any other factor which can persuade students not to cheat, therefore, ethical issues and more specifically the right of others play a more vital role in deterring them from cheating. Alongside other factors mentioned earlier, hence, the right-of-other issue turns up to serve as an influential deterring mechanism.

This study is not free from limitations. As the questionnaire was employed for collecting data, the findings are reliable to the extent that the participants answered the items honestly. In addition, as the sample was selected conveniently, they may not be representative of the population, so it is difficult to generalize the findings. Considering the significance of cheating in OAs, it is hoped that this line of research be continued by future researchers. In particular, identifying effective ways that can reduce academic dishonesty in online classes can be one area for further research. The enduring effect of "e-cheating" culture in online classes on students' perception of cheating in in-person classes can be another area for further studies.

## 7. References

- Ahanchiyan, M. R., Hoseini, M., Hashemi, F. S., & Tatari, Y. (2016). Phenomenology of cheating in exams: lived experiences of master students in Ferdowsi University of Mashhad. *Quarterly Journal of Research and Planning in Higher Education*, 22(2), 1-24.
- Ahmadi, A. (2012). Cheating on exams in the Iranian EFL context. *Journal of Academic Ethics*, 10(2), 151-170. doi: <https://doi.org/10.1007/s10805-012-9156-5>
- Amalia, R. (2018). *Students' perception of Online Assessment use in Schoology in EFL Classrooms* [UIN Sunan Ampel Surabaya].
- Bahrami, M., Hassanzade, M., Zandi, Z., Erami, E., & Kh, M. (2015). Student's attitude about cheating and its confronting strategies. *Education Strategies in Medical Sciences*, 8(2), 99-104.
- Bretag, T., Harper, R., Burton, M., Ellis, C., Newton, P., van Haeringen, K., . . . Rozenberg, P. (2019). Contract cheating and assessment design: exploring the relationship. *Assessment & Evaluation in Higher Education*, 44(5), 676-691. <https://doi.org/10.1080/02602938.2018.1527892>
- Chirumamilla, A., Sindre, G., & Nguyen-Duc, A. (2020). Cheating in e-exams and paper exams: the perceptions of engineering students and teachers in Norway. *Assessment & Evaluation in Higher Education*, 45(7), 940-957. <https://doi.org/10.1080/02602938.2020.1719975>
- Christe, B. (2003). Designing online courses to discourage dishonesty. *Educause Quarterly*, 26(4), 54-58.
- Cizek, G. (2012). *Ensuring the integrity of test scores: Shared responsibilities*. Paper presented at the annual meeting of the American Educational Research Association, Vancouver, British Columbia.
- Colnerud, G., & Rosander, M. (2009). Academic dishonesty, ethical norms and learning. *Assessment & Evaluation in Higher Education*, 34(5), 505-517. <https://doi.org/10.1080/02602930802155263>
- Curran, K., Middleton, G., & Doherty, C. (2011). Cheating in exams with technology. *International Journal of Cyber Ethics in Education (IJCEE)*, 1(2), 54-62. <https://doi.org/10.4018/ijcee.2011040105>

- Dendir, S., & Maxwell, R. S. (2020). Cheating in online courses: Evidence from online proctoring. *Computers in Human Behavior Reports*, 2, 100033. <https://doi.org/10.1016/j.chbr.2020.100033>
- Dermo, J. (2009). e-Assessment and the student learning experience: A survey of student perceptions of e-assessment. *British Journal of Educational Technology*, 40(2), 203-214. <https://doi.org/10.1111/j.1467-8535.2008.00915.x>
- Diego, L. A. B. (2017). Friends with Benefits: Causes and Effects of Learners' Cheating Practices during Examination. *IAFOR Journal of Education*, 5(2), 121-138.
- Dyer, J. M., Pettyjohn, H. C., & Saladin, S. (2020). Academic Dishonesty and Testing: How Student Beliefs and Test Settings Impact Decisions to Cheat. *Journal of the National College Testing Association*, 4(1), 1-31.
- Fishman, T. (2014). *The Fundamental Values of Academic Integrity. Second Edition (International Center for Academic Integrity)*, Available at: <https://www.academicintegrity.org/wp-content/uploads/2017/12/Fundamental-Values-2014.pdf>.
- Fluck, A. E. (2019). An international review of eExam technologies and impact. *Computers & Education*, 132, 1-15. <https://doi.org/10.1016/j.compedu.2018.12.008>
- Holden, O. L., Norris, M. E., & Kuhlmeier, V. A. (2021). Academic integrity in online assessment: A research review. *Frontiers in Education*, 6, 1-13. <https://doi.org/10.3389/educ.2021.639814>
- Howe, E. L. (2020). Perceptions of e-assessment by students and lecturers. *Cell*, 268, 76443038.
- Iskandar, N., Ganesan, N., ShafiqahEleena, N., & Maulana, A. (2021). Students' Perception Toward the Usage of Online Assessment in Universitas Putra Malaysia Amidst COVID-19 Pandemic. *Journal of Research in Humanities and Social Science*, 9(2), 09-16. <http://dx.doi.org/10.33102/alazkiyaa.v1i1.9>
- Jamil, M., Tariq, R., & Shami, P. (2012). Computer-based vs paper-based examinations: Perceptions of university teachers. *Turkish Online Journal of Educational Technology-TOJET*, 11(4), 371-381.
- Jensen, L. A., Arnett, J. J., Feldman, S. S., & Cauffman, E. (2002). It's wrong, but everybody does it: Academic dishonesty among high school and college students. *Contemporary Educational Psychology*, 27(2), 209-228. <https://doi.org/10.1006/ceps.2001.1088>
- Kennedy, K., Nowak, S., Raghuraman, R., Thomas, J., & Davis, S. F. (2000). Academic dishonesty and distance learning: Student and faculty views. *College Student Journal*, 34(2).
- Khamesan, A., & Amiri, M. (2011). The study of academic cheating among male and female students.
- Khan, S., & Khan, R. A. (2019). Online assessments: Exploring perspectives of university students. *Education and Information Technologies*, 24(1), 661-677. <https://doi.org/10.1007/s10639-018-9797-0>
- Kocdar, S., Karadeniz, A., Peytcheva-Forsyth, R., & Stoeva, V. (2018). Cheating and plagiarism in e-assessment: Students' perspectives. *Open Praxis*, 10(3), 221-235.
- Lancaster, T., & Clarke, R. (2017). Rethinking assessment by examination in the age of contract cheating.
- Ma, Y., McCabe, D. L., & Liu, R. (2013). Students' academic cheating in Chinese universities: Prevalence, influencing factors, and proposed action. *Journal of Academic Ethics*, 11(3), 169-184. <https://doi.org/10.1007/s10805-013-9186-7>
- McCabe, D. L., Trevino, L. K., & Butterfield, K. D. (2002). Honor codes and other contextual influences on academic integrity: A replication and extension to modified honor code settings. *Research in Higher Education*, 43(3), 357-378. <https://doi.org/10.1023/A:1014893102151>
- Middleton, K. V. (2020). The longer-term impact of COVID-19 on K-12 student learning and assessment. *Educational Measurement: Issues and Practice*, 39(3), 41-44. <https://doi.org/10.1111/emip.12368>

- Noorbehbahani, F., Mohammadi, A., & Aminazadeh, M. (2022). A systematic review of research on cheating in online exams from 2010 to 2021. *Education and Information Technologies*, 1-48. <https://doi.org/10.1007/s10639-022-10927-7>
- Owunwanne, D., Rustagi, N., & Dada, R. (2010). Students perceptions of cheating and plagiarism in higher institutions. *Journal of College Teaching & Learning (TLC)*, 7(11). <https://doi.org/10.19030/tlc.v7i11.253>
- Petrisor, M., Marusteri, M., Simpalean, D., Carasca, E., & Ghiga, D. (2016). Medical students' acceptance of online assessment systems. *Acta Marisiensis-Seria Medica*, 62(1), 30-32. <https://doi.org/10.1515/amma-2015-0110>
- Rogers, C. F. (2006). Faculty perceptions about e-cheating during online testing. *Journal of Computing Sciences in Colleges*, 22(2), 206-212.
- Rowe, N. C. (2004). Cheating in online student assessment: Beyond plagiarism. *Online Journal of Distance Learning Administration*, 7(2), 1-10.
- Saleh, A. M., & Meccawy, Z. (2021). EFL Female Students' Perceptions towards Cheating in Distance Learning Programmes. *English Language Teaching*, 14(1), 29-36.
- Salehi, M., & Gholampour, S. (2021). Cheating on exams: Investigating Reasons, Attitudes, and the Role of Demographic Variables. *SAGE Open*, 11(2), 21582440211004156. <https://doi.org/10.1177/21582440211004156>
- Sims, R. L. (1993). The relationship between academic dishonesty and unethical business practices. *Journal of Education for Business*, 68(4), 207-211. <https://doi.org/10.1080/08832323.1993.10117614>
- Stuber-McEwen, D., Wiseley, P., & Hoggatt, S. (2009). Point, click, and cheat: Frequency and type of academic dishonesty in the virtual classroom. *Online Journal of Distance Learning Administration*, 12(3), 1-10.
- Tuah, N. A. A., & Naing, L. (2021). Is online assessment in higher education institutions during COVID-19 pandemic reliable? *Siriraj Medical Journal*, 73(1), 61-68. <https://doi.org/10.33192/Smj.2021.09>
- Turner, S. W., & Uludag, S. (2013). Student perceptions of cheating in online and traditional classes. 2013 IEEE Frontiers in Education Conference (FIE), <https://doi.org/10.1109/FIE.2013.6685007>
- Wahid, R., & Farooq, O. (2020). Online exams in the time of COVID-19: quality parameters. *International Journal of Social Sciences and Educational Studies*, 7(4), 13-21. <http://dx.doi.org/10.23918/ijsses.v7i4p13>
- Watson, G. R., & Sottile, J. (2010). Cheating in the digital age: Do students cheat more in online courses?
- Weleschuk, A., Dyjur, P., & Kelly, P. (2019). Online assessment in higher education. *Canada: Taylor Institute for Teaching and Learning*.
- Xiong, Y., & Suen, H. K. (2018). Assessment approaches in massive open online courses: Possibilities, challenges and future directions. *International Review of Education*, 64(2), 241-263. <https://doi.org/10.1007/s11159-018-9710-5>
- Yoestara, M., Putri, Z., Keumala, M., & Idami, Z. (2020). Pre-Service English Teachers' Perception towards Online Assessment Method. *International Journal of Education, Language, and Religion*, 2(1), 1-10.
- Zhang, Y., Yin, H., & Zheng, L. (2018). Investigating academic dishonesty among Chinese undergraduate students: does gender matter? *Assessment & Evaluation in Higher Education*, 43(5), 812-826. <https://doi.org/10.1080/02602938.2017.1411467>
- Zuo, M., Ma, Y., Hu, Y., & Luo, H. (2021). K-12 students' online learning experiences during COVID-19: Lessons from China. *Frontiers of Education in China*, 16(1), 1-30. <https://doi.org/10.1007/s11516-021-0001-8>