Journal of Modern Research in English Language Studies ISSN: 2383-4242

Self-efficacy, Text Difficulty and EFL Learners' Pedagogic Task Performance

Reza Bagheri Nevisi[1*], Rasoul Mohammad Hosseinpur[2]

[1*] Assistant Professor of TEFL, University of Qom, Iran,
re.baghery@gmail.com
[2] Assistant Professor of TEFL, University of Qom, Iran,
rmhosseinpur@gmail.com

Abstract

Till recently, text difficulty has commonly been determined by employing readability formulas; however, major criticisms have been leveled against readability formulas (Graves & Graves, 2003). This research project aimed at determining text difficulty through readability formulas and Coh-Metrix. In other words, the study investigated the role of text difficulty in EFL learners' pedagogic task performances. Since both males and females took part in the study, firstly the researchers aimed at finding out whether a different pattern of task performance existed for each. Secondly, all participants were provided with two different reading passages whose difficulty levels were determined by Coh-Metrix and readability formulas. Finally, a self-efficacy questionnaire was administered to delve into learners' self-perceptions about their own performances on the pedagogic tasks. Descriptive statistics, paired samples t-test and repeated measures ANOVA, were utilized to analyze the data. The results indicated that gender of the students had no significant impact on the learners' performances on the pedagogic tasks. The findings revealed that text difficulty and the learners' self-efficacy significantly affected EFL learners' performance on the pedagogic tasks. The findings suggest that determining difficulty level of the texts through Coh-Metrix could be considered a step forward and will certainly assist language teachers and syllabus designers who strive to tailor the appropriate tasks and materials to learners at differing level of language proficiency. The results also imply that self-perceptions of learners might be a true predictor of their own performances on different tasks in general, and on pedagogic tasks in particular.

Keywords: Coh-Metrix, Pedagogic Task Types, Self-efficacy, Text Difficulty

Received 20 May 2019 Accepted 01 July 2019
Available online 01August 2019 DOI: 10.30479/jmrels.2019.10927.1368

1. Introduction

"Text difficulty" has turned into a major focus and provided a remarkable opportunity for individuals who strive to determine the suitability of a given text for a pedagogic purpose (Fulcher, 1997). Determining the difficulty level of a text accurately and precisely to the extent possible is of paramount importance to language practitioners to assure that the input to which second language readers are exposed well suits and corresponds with their processing capability and lays the foundations for the noticing, comprehension and L2 intake (Crossley, Greenfield, & McNamara, 2008). Badgett (2010) asserts that readers will probably be unable to interpret the intended meaning if the difficulty level of the texts outweighs the learners' current level of ability. The cognitive load of a text for a reader heavily relies on textual features such as lexical choice, syntactic, semantic complexity, discoursal complexity, and user's background knowledge (Sinha, Sharma, Dasgupta, & Basu, 2012). Comprehending a text does not merely hinge on within-the-text variables; reader-variables play an enormous role as well. Reader variables such as schemata, motivation, length of exposure, and prior experience have mostly gone unnoticed since these factors cannot be fully monitored and observed by the researchers (Fulcher, 1997).

Text difficulty continues to be one of the most debated, misinterpreted, and misused concepts in reading (Britton & Guelgoez, 1991; Chall, 1996; Chambliss & Calfee, 1998). It is all too commonly, but mistakenly, believed to be an exact numerical value, gained via the application of readability "formulas," which to some extent determine the difficulty level of a given text. If one adopts such an oversimplified perspective of text difficulty, then he or she might be able to say that text difficulty is a variable that resides in the text itself. Therefore, text difficulty turns into a concept that is intuitively appealing to ordinary people (Graves & Graves, 2003).

Graesser, McNamara, Louwerse, and Cai (2004) reported that recent developments in several fields have made it possible to mathematically probe into different measures of text and language comprehension that substitute superficial elements of language and rather investigate deeper-level characteristics of language. The various interdisciplinary fields that have helped achieve the above-mentioned objectives include psycholinguistics, computational linguistics, corpus linguistics, information-processing model, information retrieval, and discourse processing. All in all, the progress made in these fields has made the analysis of many deep-level textual processes automated, allowing for more precise and detailed analyses of language to happen. A number of developments in these fields have been brought about by the utilization of Coh-Metrix, a computational device which measures cohesion and text difficulty at different linguistic, discoursal, and conceptual

levels (Crossely, et al, 2008). This instrument was devised to promote reading instruction and comprehension via assisting textbook writers. It also tailors textbooks more properly and precisely to the intended audience (Graesser et al., 2004). Coh-Metrix is an advantage over traditional readability formulas such as Flesch-Kincaid and Flesch reading ease since it presents a detailed analysis of linguistic and meta-linguistic features. This tool combines semantic components, parts-of-speech, syntactic parsers, and other elements that have been developed in computational linguistics (Jurafsky & Martin, 2000). This combination makes the analysis of deeper level linguistic textual features possible. The purpose of this study is to examine if Coh-Metrix can enhance the prediction of text difficulty and compare and contrast Coh-Metrix with traditional readability formulas.

2. Literature Review

2.1. Text Difficulty

Oakland and Lane (2004) and Badgett (2010) have described readability or text difficulty as how easy or comprehensible a text is for a reader. It refers to how well a reader is able to comprehend the content of a particular text through reading. Researches have indicated that easy texts boost understanding, retention, reading pace and velocity (Sinha, et al., 2012). Generally, text difficulty means the comprehensibility and understandability of written texts (Homan, Hewit, & Linder, 1994). Berado (2006) describes text difficulty as comprising of the following elements: structural complexity, lexical density, vocabulary load, and grammatical difficulty. He also believes that text difficulty has to be thoroughly and meticulously considered when teachers and syllabus designers are going to choose appropriate materials for pedagogic purposes. Teachers and material developers take advantage of various tools and procedures to determine the readability of texts.

A compelling body of research has been carried out on text difficulty (Alexander & Jetton, 2000; Britton & Black, 1985; Britton & Guelgoez, 1991; Chall, 1996; Chambliss & Calfee, 1998; Goldman &Rakestraw, 2000; Sawyer, 1991). The fact that factors beyond those taken into account in readability formulas may have a bearing on text difficulty undoubtedly holds true. Several authors have talked about factors to consider when determining text difficulty. The best known papers on the topic are those of Anderson and Armbruster (1984) and of Beck and McKeown and their colleagues (Beck & McKeown, 1989; Beck, McKeown, & Gromoll, 1989; Beck, McKeown, & Worthy, 1995; McKeown, Beck, Sinatra, & Loxterman, 1992). Graves and Graves (2003) believe that factors are generally subdivided into two subcategories. In the first subcategory, there are six readily-defined, readily-recognized, readily-identified, and measurable factors that reside within the

text itself. However, since reading can be regarded as an interactive process that engages both the reader and the text, text factors are not totally independent of the reader. In the second subcategory, there are four factors which are less easily defined, less easily identified, and absolutely engage both the reader and the text factors which influence text difficulty and accessibility. These factors include vocabulary, sentence structure, length, elaboration, coherence and unity, text structure, familiarity of content, the required background knowledge, audience appropriateness, and quality and verve of the writing (Graves & Graves, 2003).

2.2. Readability Formulas vs. Coh-Metrix

Crossley, Allen, & McNamara (2011) regard readability formulas as one such tool to determine the difficulty level of the texts. A lot of criticisms have been leveled against these formulas because of constrained database, low dependability, ignoring higher-level features and paying attention to surface features only (i.e., syntax and vocabulary) (Crossley et al., 2011; Oakland & Lane, 2004). Ortega (2003, cited in Lu, 2008) syntactic complexity is one element in readability formulas that accounts for the degree of sophistication of syntactic structures and forms. It is one of the major factors which make a text less readable or difficult. An important factor associated with making a text syntactically difficult and more complex is sentence length which is measured in terms of average sentence length in words, number of clauses, letters, and syllables (Agnihorti & Khanna, 1992). Of course, there are a variety of other factors such as word difficulty and language structure, text structure, discourse style, genre, background knowledge, familiarity with the content, level of reasoning required, format and layout of text, and length of text which interact to influence the complexity of a particular text (Hess & Biggam, 2004). In addition, elaboration, coherence and unity, audience appropriateness, and writing quality are among other factors which influence text difficulty and accessibility (Graves & Graves, 2003).

A number of studies have examined the relationship between traditional readability formulas reading ease and L2 evaluations of readability and text difficulty (Flesch, 1948; Kincaid, Fishburne, Rogers & Chissom, 1975). These studies were undertaken because researchers were dissatisfied with traditional readability formulas when applied to texts for L2 readers. Like traditional first language readability formulas, those used in L2 have generally depended on surface-level sentence difficulty indices, such as the number of words per sentence and surface-level word difficulty indices such as syllables per words (Brown, 1998; Greenfield, 1999). Carrell (1987) discussed both the significance of developing a correct and precise L2 readability measure and the major flaws of traditional readability formulas when applied to L2 texts. She maintained that more accurate readability

formulas were necessary to guarantee a good correspondence between L2 reading texts and L2 learners. She criticized traditional readability formulas for not taking care of reader characteristics or for within-the-text elements such as syntactic complexity, rhetorical organization, and propositional density. Brown (1998) stated that traditional readability formulas couldn't account for L2 reader-based variables. In addition, he mentioned that readability formulas for L2 readers should take the type, function, and frequency of words and word redundancy into consideration.

Crossely et al. (2008) state that Coh-Metrix seems to be advantageous over readability formulas since it takes linguistic and meta-linguistic features of a text into account. They point to the fact that Coh-Metrix takes care of the established links between features of the text and stored mental representations of readers. These representations are not merely linguistic but involve world knowledge, knowledge of text genre, and the discourse model.

Evidently, determining the text difficulty through Coh-Metrix is more effective for a number of reasons. First, it is a psycholinguistically-based evaluation of text difficulty that goes beyond surface readability features. Second, the readers' interaction with a text is one major component that traditional readability formulas fail to account for, and through Coh-Metrix, attempts are made to rectify this problem. Third, it incorporates measures of text cohesion and meaning construction into consideration as well (Gernsbacher, 1997; McNamara et al., 1996). Forth, it encodes meaning as a multi-stage and multi-dimensional process (Koda, 2005). This encoding could involve measures such as decoding, syntactic parsing, and meaning construction (Just & Carpenter, 1987; Perfetti, 1985; Rayner & Pollatsek, 1994). Fifth, it is a readability measure that takes proper account of the role of short-term memory and the restrictions it imposes (Crossely, et al, 2008).

Considerably, little attention has been given to the empirical validation of traditional readability formulas in relation to L2 contexts. Even less has been given to developing alternatives more in line with current knowledge about psycholinguistic models of L1 or L2 reading. Most, if not all, studies that have investigated readability formulas for L2 students have depended on traditional readability measures (e.g., Brown, 1998; Greenfield, 1999, 2004; Hamsik, 1984). Therefore, the present study sets to answer the following research questions.

1. Will gender play a role in learners' pedagogic task performances?
2. Will self-efficient learners perform differently compared to their non-efficient counterparts with regard to their pedagogic task performances?
3. Will text difficulty play a role in learners' pedagogic task performances?

4. In what ways is using Coh-Metrix advantageous over using readability formulas to determine the difficulty level of texts?

3. Method

3.1. Participants

Ninety Iranian undergraduate students took part in the study. They were all taking a reading comprehension course at the time. Thirty-four male and fifty-six female participants of two different reading proficiency levels i.e. high and low were included in the research project. The participants were adults and of Persian and Turkish language background. Due to the administrative difficulties of randomization, available sampling was utilized. Available reading comprehension classes were taken advantage of.

3.2. Procedure

3.2.1. Phase 1

The first methodological step to be taken was to place EFL learners who were all taking reading comprehension (3) at the time into two (high and low) different levels of language proficiency. To achieve this objective, a standardized TOEFL test was administered. Based on the results of the test, the participants were placed into two different levels of language proficiency. Those who scored 14.71 and below formed the low proficiency group and the participants whose scores fell above 14.72 made up the high proficiency group.

3.2.2. Phase 2

Having determined the language proficiency level of EFL learners, the researchers administered two different reading passages with three different pedagogic task types (a comprehension-check task, a restatement task, and a synonym task). Two reading passages were included in the research project to account for the text difficulty variable. The first passage was extracted from (Nilipour, 1996) and the second one was borrowed from (Marerlli & Nadler, 1989). The first text entitled (We think with our muscles) was the simpler text and the second text *(*The Indian character) was a more difficult one based on readability statistics.

3.2.3. Phase 3

The self-efficacy questionnaire (Kitikanan & Sasimonton, 2017) was administered when the learners had completed their performances on the pedagogic tasks. The self-efficacy questionnaire was administered immediately after learners' completing pedagogic task types of the two texts. The self-efficacy questionnaire helped the researchers find out the learners' self-

perceptions about their own performances across different pedagogic task types. Once participants had completed the pedagogic tasks, they were provided with the self-efficacy questionnaire. The questionnaire asked them to rate their performances on each pedagogic task in particular, and their overall performance in general. This would help the researchers to delve into their self-perception of their own performances and to see to what extent the findings would match and go with their actual performances.

Table 1

Readability Statistics

Readability statistics criteria	We think with our muscles	The Indian character
Words	201	632
Sentences	14	29
Sentence per paragraph	3.5	7.2
Words per sentence	14.35	21.7
Passive sentences	28%	24%
Flesch reading Ease	32.2	61.8
Flesh-Kincaid Grade level	4.2	8.4

3.4. Pedagogic Tasks

Three different pedagogic tasks were developed for each text. A comprehension-check task, a restatement task, and a synonym task were included for each text. As for the comprehension-check task, the participants were asked to answer the questions based on the provided text. Participants were asked to recognize and produce the answer for the comprehension-check task from within the text. As for the restatement task, some sentences of the texts were chosen to be paraphrased. The participants were asked to restate the sentences in a more simplified manner. They were asked to comprehend, process, produce, and manipulate the necessary pieces of information all at once. This time, their production was not limited as they were asked to substitute the sentences with their own words. As for the synonym task, the participants were provided with five sentences and one word in italics for which they were asked to choose one synonym from among five possible options.

3.5. Task Complexity of the Pedagogic Tasks

The complexity of each pedagogic task was to be determined in advance. This was achieved based on the existing models of task complexity in the literature; i.e. Brown (1984), Prabhu (1987), Skehan (2001), and Robinson (2001).

Based on the Brown's model of task complexity (1984), tasks are classified into static, dynamic, and abstract ones with abstract being the most difficult type. Based on his model, the comprehension-check, the restatement task, and the synonym task were dynamic, abstract, and static respectively.

Based on the Prabhu's model of task complexity (1987), the following factors were taken into account to determine the complexity of the pedagogic tasks. (1) amount of information (AI), amount of reasoning (AR), (3) degree of precision (DP), (4) degree of familiarity (DF), and (5) degree of abstractness(DA).

Table 2

Prabhu's Model of Task Complexity

	AI	AR	DP	DF	DA
Synonym	Lowest	Low	Highest	High	Low
Comprehension-check	High	Average	High	High	Low
Restatement	High	High	Low	Low	High

As for synonym task, fewer elements were transferred in terms of the amount of information needed, in comparison with the other two pedagogic tasks, namely, comprehension-check and restatement tasks. In terms of the amount of reasoning needed to complete the tasks, the restatement task needed the most and the synonym task needed the least. The amount of reasoning has to do with the number of steps that need to be taken for carrying out a task. For instance, while doing the synonym task, the participants were asked to match the words in italics with the one appropriate answer from among four possible options. With regard to the degree of precision needed for each pedagogic task, the synonym task was the most precise while the restatement task needed less precision in terms of the answer provided by the participants. The answer for the comprehension-check needed to be more precise than that of the restatement task. As for the restatement task, participants were asked to produce the answer of their own, but the response to the comprehension-check task was to be spotted and merely written from within the text. In terms of the familiarity, the restatement task was the least familiar to the participants while the other two pedagogic tasks were more familiar to them.

Based on the Skehan's model of task complexity (2001), the following factors were taken into account: (1) code complexity, (2) cognitive complexity, (3) communicative stress, and (4) learner factors.

Table 3

Skehan's Model of Task Complexity

	Code complexity	Cognitive complexity	Communicative stress	Learner factors
Synonym	Low	Low	Low	Low
Comprehension-check	Average	Average	Average	Average
Restatement	High	High	High	High

In terms of code complexity, the restatement task needed a more linguistically-complex and various response on the part of language learners while the synonym task required the least and comprehension-check task stood somewhere in the middle. In terms of the cognitive complexity, the restatement task was the most complex one, due to the degree of familiarity of discourse, genre, and task. On the contrary, the synonym task was the least cognitively complex task because the participants merely matched a word with its synonym and no memory load was imposed on them. In terms of the communicative stress, the restatement task required more time to be completed and the responses were more open-ended. Finally, with regard to the learner factors, the restatement task brought their intelligence, imagination and their personal experiences to close touch with the reality. Robinson's model of task complexity (2001) takes into account the following factors: (1) cognitive factors, (2) interactive factors, and (3) difficulty factors.

Table 4

Robinson's Model of Task Complexity

	Cognitive	Interactive	Difficulty
Synonym	Low	Low	Low
Comprehension-check	Average	Average	Average
Restatement	High	High	High

Based on the Robinson's model of task complexity, the restatement task was the most complex one since it required the learners to plan ahead of time, to tap on their prior knowledge, and also apply many elements in comparison with the other two pedagogic tasks which required less planning, fewer elements and also less activation of the prior knowledge on the part of EFL learners. In terms of the interactive factors, the response for the synonym task was closed but that of the restatement task was more open-ended. With regard to difficulty factors, the restatement task created more anxiety on the part of language learners and required more confidence and proficiency on their part.

4. Results and Discussion

4.1. Results

4.1.1. Gender and Pedagogic Task Performance

The results of descriptive statistics on gender show that gender did not have a significant impact on learners' pedagogic task performance as the mean score for males and females stood at 15.99 and 15.74 respectively. Table 5 shows the descriptive statistics on gender.

Table 5

Descriptive Statistics: Gender on Pedagogic Tasks

			95% Confidence Interval	
Gender	Mean	Std. Error	Lower Bound	Upper Bound
Male	15.992	.335	15.318	16.666
Female	15.743	.209	15.322	16.163

A repeated-measures ANOVA was run to compare the mean scores of the male and female students on the Comprehension-check, Restatement and Synonym tasks.

Table 6

Repeated Measures ANOVA for Gender Effect

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Gender	3.758	1	3.758	.398	.531	.008
Error	453.368	48	9.445			

The F-observed value for the effect of the gender of the students was .39 (Table 6). This amount of F-value is lower than the critical value of F at 1 and 48 degrees of freedom, that is to say, 4.04. An effect size (partial eta squared) of .008 is considered to be of no significance at all. The results indicate that the gender of the students did not have any significant effect on their performance on the pedagogic tasks.

The overall mean score for females was 15.73 while the mean score for males was 15.99. Repeated Measures ANOVA was run to see whether that amount of difference between the mean performance of the two groups was statistically significant or not. The results of the repeated Measures ANOVA indicated that F-observed was lower than F-critical at 1.48 degrees of freedom, but Partial Eta Squared indicated that the effect size was weak. The gender of the students did not have a significant impact on the learners' performances on the pedagogic tasks.

Table 7

Difficulty Level, Task Types and Students' Gender

Gender	Tasks	LEVEL	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Male	Comprehension-Check	Easy	19.286	.291	18.701	19.870
		Difficult	18.571	.590	17.385	19.758
	Restatement	Easy	11.746	.731	10.276	13.216
		Difficult	6.349	.840	4.661	8.037
	Synonyms	Easy	20.000	.152	19.694	20.306
		Difficult	20.000	.000	20.000	20.000
Female	Comprehension-Check	Easy	19.722	.181	19.358	20.087
		Difficult	17.685	.368	16.945	18.425
	Restatement	Easy	10.988	.456	10.071	11.904
		Difficult	6.173	.524	5.120	7.226
	Synonyms	Easy	19.889	.095	19.698	20.080
		Difficult	20.000	.000	20.000	20.000

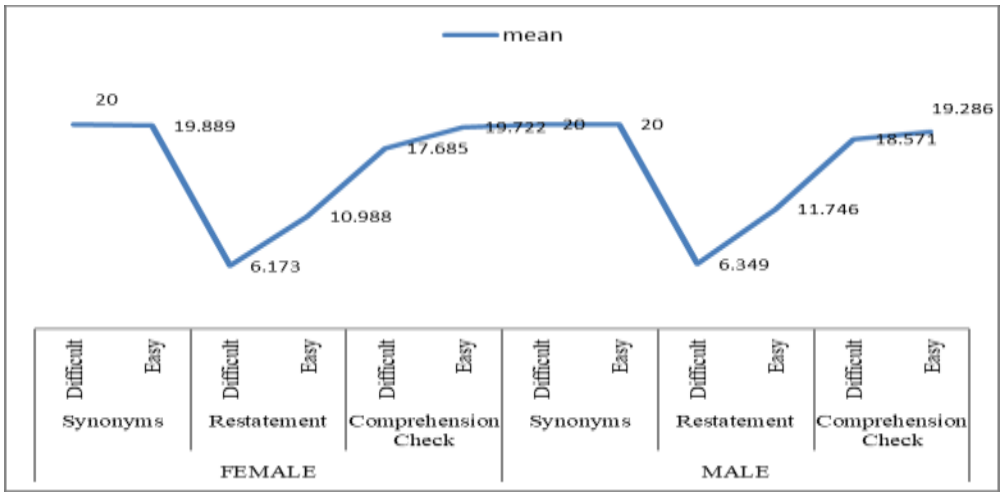

Figure 1. Difficulty Level, Task Types and Students' Gender

4.1.2. Self-Efficacy and Pedagogic Task Performance

As displayed in Table 8, the overall mean scores for the self-efficient and non-self-efficient learners were16.16 and 14.69 respectively.

Table 8

Descriptive Statistics; Self-Efficacy on Pedagogic Tasks

			95% Confidence Interval	
Self-Efficacy	Mean	Std. Error	Lower Bound	Upper Bound
Self-efficient	16.165	.176	15.811	16.519
Non self-efficient	14.698	.313	14.068	15.328

Table 9

Repeated-Measures ANOVA, Task Types, Self-Efficacy and Difficulty Level

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Tasks	Pillai's Trace	.953	476.495[a]	2.000	47.000	.000	.953
	Wilks' Lambda	.047	476.495[a]	2.000	47.000	.000	.953
	Hotelling's Trace	20.276	476.495[a]	2.000	47.000	.000	.953
	Roy's Largest Root	20.276	476.495[a]	2.000	47.000	.000	.953
Tasks* Self-efficacy	Pillai's Trace	.212	6.328[a]	2.000	47.000	.004	.212
	Wilks' Lambda	.788	6.328[a]	2.000	47.000	.004	.212
	Hotelling's Trace	.269	6.328[a]	2.000	47.000	.004	.212
	Roy's Largest Root	.269	6.328[a]	2.000	47.000	.004	.212
Level	Pillai's Trace	.849	269.879[a]	1.000	48.000	.000	.849
	Wilks' Lambda	.151	269.879[a]	1.000	48.000	.000	.849
	Hotelling's Trace	5.622	269.879[a]	1.000	48.000	.000	.849
	Roy's Largest Root	5.622	269.879[a]	1.000	48.000	.000	.849
Level * Self-efficacy	Pillai's Trace	.112	6.055[a]	1.000	48.000	.018	.112
	Wilks' Lambda	.888	6.055[a]	1.000	48.000	.018	.112
	Hotelling's Trace	.126	6.055[a]	1.000	48.000	.018	.112
	Roy's Largest Root	.126	6.055[a]	1.000	48.000	.018	.112
Tasks * Level	Pillai's Trace	.852	134.765[a]	2.000	47.000	.000	.852
	Wilks' Lambda	.148	134.765[a]	2.000	47.000	.000	.852
	Hotelling's Trace	5.735	134.765[a]	2.000	47.000	.000	.852
	Roy's Largest Root	5.735	134.765[a]	2.000	47.000	.000	.852
Tasks * Level * Self-efficacy	Pillai's Trace	.217	6.525[a]	2.000	47.000	.003	.217
	Wilks' Lambda	.783	6.525[a]	2.000	47.000	.003	.217
	Hotelling's Trace	.278	6.525[a]	2.000	47.000	.003	.217
	Roy's Largest Root	.278	6.525[a]	2.000	47.000	.003	.217

A repeated-measures ANOVA was run to compare the mean scores of the self-efficient and non-self-efficient students on the Comprehension-Check, Restatement and Synonym tasks. As table 10 indicates, there is a significant interaction between task types and students' Self-Efficacy (F (2.47) = 6.32; P = .004). In the same vein, there is a significant interaction between difficulty level and students' Self-Efficacy (F (1.48) = 6.05; P = .47). Task types and difficulty level of (F (2.47) = 134.76; P = .000) also interacted significantly. The interaction between difficulty level, task types and students' Self-Efficacy was significant as well (F (2.47) = 6.52; P = .003).

The overall mean score for efficient learners was 16.16 while the mean score for non-efficient was 14.69. Repeated Measures ANOVA was run to discover whether that amount of difference between the mean performance of the two groups was statistically significant or not. The results of repeated Measures ANOVA indicated that F-observed was higher than F-critical at 1.48 degrees of freedom and Partial Eta Squared indicated that the effect size was strong. The self-efficacy of the students had a significant effect on the learners' performance on the pedagogic tasks. Thus, the null-hypothesis as learners' self-efficacy will not play a role in their performance on pedagogic tasks is rejected.

Based on these results it could be concluded that the Self-Efficacy of the students had a significant effect on their performances on the pedagogic tasks.

The F-observed value for the effect of the Self-Efficacy of the students was 16.66 (Table 10). This amount of F value is higher than the critical value of F at 1.48 df.

Table 10

Repeated Measures ANOVA for Self-Efficacy Effect

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared
Self-efficacy	117.792	1	117.792	16.662	.000	.258
Error	339.334	48	7.069			

The results indicated that the Self-Efficacy of the students had a significant effect on their performances on the pedagogic tasks; however, the results must be interpreted with great care F(1.48) = 16.66 p = 0.000).

Table 11 displays the descriptive statistics for difficulty level, task types and students' Self-Efficacy. With regard to self-efficacy level, high-efficient learners performed much better on the pedagogic task types in comparison with their low-efficient counterparts.

Table 11

Difficulty Level, Task Types and Students' Self-Efficacy

Self-efficacy	Tasks	Level	Mean	Std. Error	95% Confidence Interval	
					Lower Bound	Upper Bound
Self-efficient	Comprehension Check	Easy	19.649	.179	19.290	20.009
		Difficult	18.509	.322	17.861	19.157
	Restatement	Easy	11.871	.401	11.065	12.678
		Difficult	6.959	.461	6.032	7.887
	Synonyms	Easy	20.000	.090	19.820	20.180
		Difficult	20.000	.000	20.000	20.000
Quite Self-efficient	Comprehension Check	Easy	19.444	.318	18.805	20.084
		Difficult	16.111	.574	14.958	17.265
	Restatement	Easy	9.074	.714	7.638	10.510
		Difficult	3.889	.821	2.238	5.539
	Synonyms	Easy	19.667	.160	19.346	19.988
		Difficult	20.000	.000	20.000	20.000

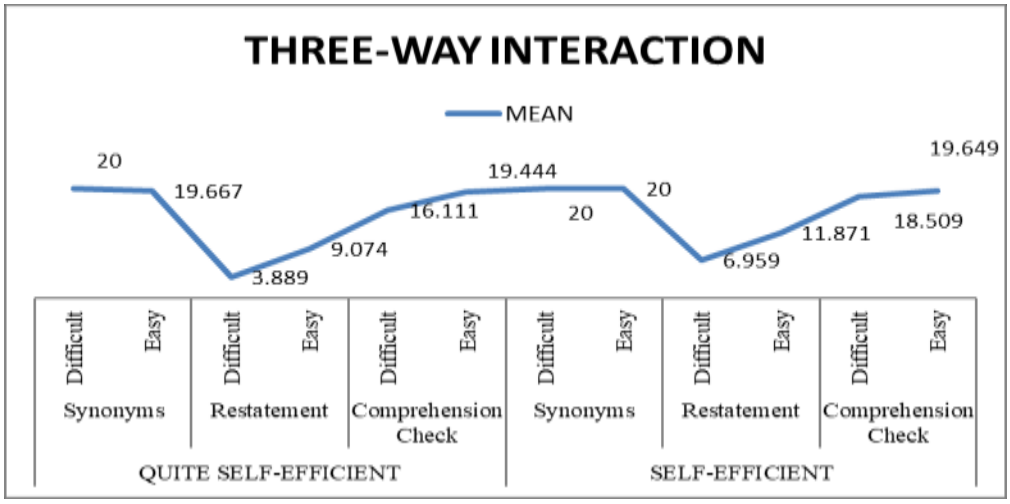

Figure 2. Difficulty Level, Task Types and Students' Self-Efficacy

4.1.3. Text Difficulty and Pedagogic Task Performance

As displayed in Table 12, the participants performed better on the easy text with a mean of 14.37 while their average mean on the difficult text was 10.77.

Table 12

Descriptive Statistics for Text Difficulty on Pedagogic Tasks

	Mean	N	Std. Deviation	Std. Error Mean
Total Easy	14.37	90	2.329	.246
Total Difficult	10.77	90	3.233	.341

A paired-samples t-test was run to compare the mean scores of the students on the total easy and difficult tasks to probe whether text difficulty played any role in EFL readers' performances on different pedagogic tasks. The t-observed value was 25.66 (Table 13). This amount of t-value is higher than the critical value of 1.98 at 89 degrees of freedom.

Table 13

Paired-Samples t-test; Text Difficulty on Pedagogic Tasks

Paired Differences								
			95% Confidence Interval of the Difference					
Mean	Std. Deviation	Std. Error Mean	Lower	Upper	T	df	Sig. (2-tailed)	
3.600	1.331	.140	3.321	3.879	25.665	89	.000	

Since the t-observed value is higher than its critical value, it can be said that the difference between the total mean scores of easy and difficult texts was significant.

4.1.4. Using Coh-Metrix to Assess Cohesion and Text Difficulty

The two different texts already mentioned in the method section were analyzed using Coh-Metrix indices of cohesion and difficulty: Argument overlap, latent semantic analysis (LSA), and the number of connectives to assess both cohesion and difficulty. The results were very much detailed compared to the ones obtained from readability formulas (Flesch-Kincaid Grade Level, Klare, 1974, 1975).

The LSA score for each text was calculated by calculating the average cosine between each sentence in the section and the section as a whole. Higher LSA cosines imply higher cohesion. For the simpler texts, the results relatively reflected the findings for FKGL showing a slight decrease in cohesion (indicating an increase in difficulty for this text) across sentences. The more difficult text received the highest LSA values. Argument Overlap for both texts approximately corresponded. As can be seen from Table 14, both simpler and more difficult texts had the same argument overlap. This was in line with our prediction that both texts would have similar cohesion levels. As for the number of connectives, the number varied and the more

difficult text (*the Indian character*) enjoyed more connectives compared to the simpler text (*We think with our muscles*).

Table 14

Using Coh-Metrix to Assess Cohesion and Text Difficulty

Coh-Metrix	We think with our muscles	The Indian character
Argument overlap	8	10
Latent semantic analysis	14	29
Number of connectives	11	26
Coh-Metrix index	3.2	7.4

4.2. Discussion

This study was designed to explore the impact of gender, text difficulty, and self-efficacy of EFL language learners on their pedagogic task performances. Firstly, comparing and contrasting the performances of males and female EFL learners on the pedagogic task types was one of the objectives of the study. Secondly, the researchers sought to delve into the learners' self-perceptions of their own pedagogic task performances and investigated the role of self-efficacy in learners' pedagogic task performance. Thirdly, the role of text difficulty in EFL learners' pedagogic task types was explored. Two different instruments were utilized to determine the difficulty level of texts (readability formula and Co-Metrix). The relative effectiveness of both devices was explored.

Two reading passages were included in the research project to account for the text difficulty variable. The second text also was a more difficult text based on the score obtained from the readability statistics and Coh-Metrix. The participants performed differently on the two texts and text difficulty had a significant impact on learners' pedagogic task performance. Coh-Metrix was a more effective means of determining text difficulty because it went beyond surface features of the text and took coherence and cohesion into account as well. Apparently, Coh-Metrix is a more effective means of determining the text difficulty. One possible explanation for the relative effectiveness of using Coh-Metrix to determine text difficulty in comparison with the more traditional methods is that Coh-Metrix is a psycholinguistically-oriented assessment of text difficulty that goes beyond surface readability features. Another justification would be that it attempts to take the readers' interaction with a text into account. It incorporates measures of text cohesion and meaning construction into consideration as well (Gernsbacher, 1997; McNamara et al., 1996) and it encodes meaning as a multi-stage and multi-dimensional process (Koda, 2005). Last but not necessarily the least logic behind using Coh-Metrix as a more appropriate device to determine the text difficulty is that it considers the role of short-term memory and its imposed limitations more properly (Crossely, et al, 2008).

The overall findings of this study, aligned with the findings of many other studies (e.g., Alexander & Jetton, 2000; Britton & Black, 1985; Britton & Guelgoez, 1991; Chall, 1996; Chambliss & Calfee, 1998; Goldman & Rakestraw, 2000; Linderholm, et al., 2000; Muth, 1989; Greenfield, 2004; Tamor,1981) support the utilization of Coh-Metrix to determine the text difficulty suggesting that it is apsycholinguistically-based model of text comprehensibility that goes beyond the surface features that a readability formula takes care of. The comparison of the two instruments revealed that Coh-Metrix was more effective in determining the text difficulty. This finding concurs with the results of previous studies (e.g. Crossley et al., 2011; Oakland & Lane, 2004). Traditional readability formulas, however, are generally not supported by or founded upon theories of reading or comprehension, but on tracing statistical relations (Crossely, et al, 2008). One possible shortcoming of readability formulas is their weak construct validity. The restricted validity of these formulas has prompted several researchers within the field of discourse processing to apply them cautiously (Davison & Kantor, 1982; Rubin, 1985). According to Crossley et al. (2008) such formulas are being commonly used for a variety of purposes for different texts and readers because of their simplicity and objectivity. They also assert that shortcomings of traditional formulas become more obvious when one compares and contrasts them with psycholinguistic models like Coh-Metrix that account for cohesion and coherence as well. Psycholinguists see reading as a multi-dimensional skill acting at various stages of processing: lexical, syntactic, semantic, and discoursal (Just & Carpenter, 1987; Koda, 2005).

Another concept the researchers delved into was whether self-efficient learners could outperform their non-self-efficient counterparts when performing pedagogic tasks. It came to light that self-efficient learners had better self-perceptions about their own pedagogic task performances. This might partly be due to their proficiency level which in turn led to better self-confidence and finally resulted in a much better performance on the pedagogic task types.

5. Conclusion and Implications

The present study aimed at finding out whether male participants performed differently on pedagogic task types in comparison with their female counterparts. As results indicated, there was no significant difference between males and females with regard to their performances on the pedagogic task types. Therefore, gender did not play any role in learners' pedagogic task performances.

Once students had completed their performances on the pedagogic tasks, they were provided with the self-efficacy questionnaire. The questionnaire asked students to rate their performances on each pedagogic

task in particular, and their overall performance in general. This would help the researches to delve into learners' self-perception of their own performances and to see whether it matched their actual performances on the very same pedagogic tasks. Based on these results, it could be concluded that the learners' self-efficacy significantly impacted their pedagogic task performances. It can also be implied that EFL learners with higher self-efficiency can outperform those with lower self-efficiency when they are carrying out tasks in general and pedagogic task in particular.

The study might add to the existing literature on text difficulty specification through taking into considerations the Coh-Metrix which is a psycholoinguistically-oriented tool to determine text difficulty. Furthermore, the self-efficacy questionnaire administered in this study can provide invaluable insights into how students' perceptions on their performances can have any possible bearings on their future task completion in classroom settings. Another important point to bear in mind is that the models of task complexity presented in this study can shed further light on the way both pedagogic and target tasks can be best designed, implemented, and evaluated. Interested researchers within the field of applied linguistics are also recommended to carry more in-depth research on task types and text difficulty in general, and pedagogic task types, and Coh-Metrix in particular to delve into various aspects of such variables in both EFL and ESL settings. In the same vein, Coh-Metrix can be used to assess textual cohesion in textbooks and can deliver more than 300 indices of textual cohesion and difficulty (McCarthy, Lightman, Dulfty, &McNamara, 2019; Graesser, McNamara, & Kulikovich, 2011).

The results of the present research will possibly have a number of implications for curriculum developers, English teachers, and learners. Curriculum developers may benefit from the findings of this study when engaged in designing pedagogic tasks and determining text difficulty and will have to consider the importance of task types, task design, task and text complexity, and task evaluation. English teachers could be familiarized with the ways to design task types with varying complexity levels to serve different purposes. Language learners will also be provided with a better and a more vivid picture of task complexity and text difficulty, with an emphasis on Coh-Metrix. Through raising the learners' consciousness about the role of traditional and more recent readability formulas, language learners will be better informed and cognizant about the optimal amount of time and effort needed for mastering different pedagogic tasks and comprehending text types.

The study has delimited itself to three specific pedagogic task types: Comprehension-check, synonym, and restatement. Another potential delimitation of the study could be the self-efficacy questionnaire (Kitikanan

& Sasimonton, 2017) used in this study since a different questionnaire could have been utilized. However, this specific one well matched the already-set purposes of the study. Finally, the researchers limited themselves to the three well-established models of task complexity in the literature: Those of Prabhu (1987), Skehan (2001), and Robinson (2001).

References

Agnihorti, R.K., & Khanna, A.L. (1992). Evaluating the readability of school text books: An Indian study. *Journal of Reading, 35*(4), 282-288.

Alexander, P. A, & Jetton, T. L. (2000). Learning from text: A multidimensional and developmental perspective. In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of reading research, volume III* (pp. 285-310). Mahwah, NJ: Erlbaum.

Anderson, T. H., & Armbruster, B. B. (1984). Content area textbooks. In R. C. Anderson, J. Osborn, & R. J. Tierney (Eds.), *Learning to Read in American Schools* (pp. 193-226). Hillsdale, NJ: Erlbaum.

Badgett, B.A. (2010). Toward the development of a model to estimate the readability of credentialing-examination materials. UNLV, Paper 185. Beck, I. L., & McKeown, M. G. (1989). Expository text for young readers. The issue of coherence. In L. Resnick (Ed.), *Essays in honor of Robert Glaser* (pp. 47-66), Hillsdale, NJ: Erlbaum.

Beck, I. L., McKeown, M. G., & Gromoll, E. W. (1989). Issues that may affect social studies learning: Examples from four commercial programs. *Cognition and Instruction*, *6*(1), 99-158.

Beck, I. L., McKeown, M. G., & Worthy, J. (1995). Giving a text voice can improve students' understanding. *Reading Research Quarterly*, *30*(2), 220-238.

Britton, B. K., & Black, J. B. (Eds.). (1985). Understanding expository text: A theoretical and practical handbook for analyzing explanatory text. Hillsdale, NJ: Erlbaum

Britton, B. K; Guelgoez, S. (1991). Using Kintsch's computational model to improve instructional text: Effects of repairing inference calls on recall and cognitive structures. *Journal of Educational Psycholog*y, *83*(3), 329-345.

Berardo, S. A. (2006). The use of authentic materials in the teaching of reading. *The Reading Matrix,6*(2), 60-68.

Brown, J. D. (1998). An EFL readability index. *JALT Journal, 29*(2), 7–36.

Carrell, P.L. (1987). Readability in ESL. *Reading in a Foreign Language, 4*(1), 21-40.

Chall, J. S., Bissex, G. L., Conard, S. S., & Harris-Sharples, S. (1996). *Qualitative assessment of text difficulty: A practical guide for teachers and writers*. Cambridge, MA: Brookline Books.

Chall, J., & Dale, E. (1995). *Readability revisited: The new Dale-Chall readability formula*. Cambridge, MA: Brookline Books.

Chambliss, M., & Calfee, R. C. (1998). *Textbooks for learning: Nurturing children's minds*. Cambridge, MA: Basil Blackwell.

Crossley, S.A., Louwerse, M. M., McCarthy, P.M., & McNamara, D. (2007). A linguistic analysis of simplified and authentic texts. *The Modern Language Journal*, *91*(1), 15-30.

Crossley, S.A., Greenfield, J., & McNamara, D.S. (2008). Assessing text readability using cognitively based indices. *TESOL Quarterly, 42*(3), 475-493.

Crossley, S.A., & McNamara, D.S. (2008). Assessing L2 reading text at the intermediate level: An approximate replication of Crossley, Louwerse, McCarthy, & Mcnamara (2007). *Language Teaching, 41*(3), 409-429.

Crossley, S. A., Allen, D.B., & McNamara, D.S. (2011). Text readability and intuitive simplification: A comparison of readability formulas. *Reading in a Foreign Language, 23*(1), 84-101.

Davison, A., & Kantor, R. (1982). On the failure of readability formulas to define readable texts: A case study from adaptations. *Reading Research Quarterly, 17(2),* 187–209.

Flesch, R. (1948). A new readability yardstick. *The Journal of Applied Psychology, 32,*221–233.

Fulcher, G. (1997). Text difficulty and accessibility: Reading formulae and expert judgment. *System, 25*(4), 404-519.

Gernsbacher, M. (1997). Coherence cues mapping during comprehension. In J. Costermans & M. Fayol (Eds.), *Processing inter clausal relationships. Studies in the production and comprehension of text* (pp. 3–22). Mahwah NJ: Erlbaum.

Goldman, S. R., & Rakestraw, J. A., Jr. (2000). In M. Kamil, P. Mosenthal, P. D. Pearson, & R. Barr (Eds.), *Handbook of Reading Research, volume III* (pp. 3-11-335). Mahwah, NJ: Erlbaum.

Graesser, A. C., McNamara, D. D., Louwerse, M. L., & Cai, Z. (2004). Coh-Metrix: Analysis of text on cohesion and language. *Behavior Research Methods, Instruments, & Computers, 36,* 193–202.

Graesser, A. C., McNamara, S. D., & Kulikovich, M. J. (2011). Coh-Metrix: Providing multilevel analyses of text characteristics. *Educational Researcher, 40*(5), 223-234.

Graves, M.F., & Graves, B.B. (2003). *Scaffolding reading experiences: Designs for student success.* Christopher-Gordon.

Greenfield, G. (1999). *Classic readability formulas in an EFL context: Are they valid for Japanese speakers?* Unpublished doctoral dissertation, Temple University, Philadelphia,PA, United States.

Greenfield, J. (2004). Readability formulas for EFL.*JALT Journal, 26*(1)*,* 5–24.

Haberlandt, K. F., & Graesser, A. C. (1985). Component processes in text comprehension and some of their interactions. *Journal of Experimental Psychology. General, 114,* 357–374.

Hamsik, M. J. (1984). *Reading, readability, and the ESL reader.* Unpublished doctoral dissertation, University of South Florida.

Hess, K. & Biggam, S. (2004). A discussion of increasing text complexity. Appendices for New England Common Assessment Program (NECAP).

Homan, S., Hewitt, M., Linder, S. (1994). The development and validation of a formula for measuring single-sentence test item readability. *Journal of Educational Measurement, 31*, 349-358.

Jurafsky, D., & Martin, J. H. (2000). *Speech and language processing: An introduction to natural language processing, computational linguistics, and speech recognition*. Upper Saddle River, NJ: Prentice-Hall.

Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review, 87(*4), 329–354.

Just, M. A., & Carpenter, P. A. (1987). *The psychology of reading and language comprehension*. Boston: Allyn & Bacon, Inc.

Kincaid, J. P., Fishburne, R. P., Rogers, R. L., & Chissom, B. S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel*, Research Branch Report 8–75, Millington,TN: Naval Technical Training, U. S. Naval Air Station, Memphis, TN.

Kitikanan, P. & Sasimonton, P. (2017). The relationship between English self-efficacy and English learning achievement of L2 Thai learners. *Language Education and Acquisition Research Network Journal, 10*(1),148-163.

Koda, K. (2005). *Insights into second language reading*. Cambridge: Cambridge University Press.

Linderholm, T., Everson, M. G., van den Broek, P., Mischinski, M., Crittenden, A., & Samuels, J. (2000). Effects of causal text revisions on more- and less-skilled readers' comprehension of easy and difficult texts. *Cognition & Instruction*, *18*, 525-556.

Lu, X. (2008). Automatic measurement of syntactic complexity using the revised developmental level scale.American Association for Artificial Intelligence.

McCarthy, M. P., Lightman, J. E., Dufty, F. D., & McNamara, S. D. (2019). Using Coh-Metrix to assess cohesion and difficulty in high-school textbooks.

McKeown, M. G., Beck, I. L., Sinatra, G. M., & Loxterman, J. A. (1992). The contribution of prior knowledge and coherent text to comprehension. *Reading Research Quarterly,27*, 78-93.

McNamara, D. S., Kintsch, E., Butler-Songer, N., & Kintsch, W. (1996). Are good texts always better? Interactions of text coherence, background knowledge, and levels of understanding in learning from text. *Cognition and Instruction, 14,* 1–43.

McNamara, D. S., Louwerse, M. M., & Graesser, A. C. (2002). Coh-Metrix (Version 2.0) [Software]. Memphis, TN: University of Memphis, Institute for Intelligent Systems. Available from http://cohmetrix.memphis.edu/cohmetrixpr/index.html

Oakland, T. & Lane, H. (2004). Language, reading, and readability formulas: Implications for developing and adapting tests. *International Journal of Testing 4*, 239-252.

Perfetti, C. A. (1985). *Reading ability*. Oxford: Oxford University Press.

Rayner, K., &Pollatsek, A. (1994). *The psychology of reading*. Englewood Cliffs, NJ: Prentice Hall.

Prabhu, S. N. (1987). *Second language pedagogy*. Oxford: Oxford University Press.

Robinson, P. (2001). Task complexity, task difficulty, and task production: exploring interactions in a componential framework. *Applied Linguistics, 22*(1), 27-57.

Rubin, A. (1985). How useful are readability formulas? In J. Osborn, P. T. Wilson, & R. C. Anderson (Eds.), *Reading education: Foundations for a literate America* (pp. 61–77). Lexington, MA: Lexington Books.

Sawyer, M. H. (1991). A review of the research in revising instructional text.*Journal of Reading Behavior*, *23*(3), 307-333.

Sinha, M., Sharma, S., Dasgupta, T., &Basu, A. (2012). New readability measure for Bangla and Hindi texts. *Proceedings of COLING*, Mumbai.

Skehan, P. (2001). Tasks and language performance assessment. In M. Bygate et al. (Eds). *Researching pedagogic tasks. Second language learning and testing* (pp.99-118). Harlow: Longman.

Tamor, L. (1981). Subjective text difficulty: an alternative approach to defining the difficulty level of written text. *Journal of Reading Behavior*, *XIII* (2), 165-172.