Readability Index and Reading Complexity in high school EFL Textbooks

Mohammad Abdollahi-Guilani *
(corresponding author)
Assistant Professor of English at Qazvin International University, Buin Zahra Center for Higher Technical and Engineering Education, Qazvin, Iran
Email: abdollahi20@gmail.com

ABSTRACT
This study aims to investigate the readability of Iranian English textbooks, Vision series 1-3, for high school. It focused on the reading sections of student books and workbooks. The texts were collected and fed into different readability tools such as Coh-metrix, and Flasch-Kincaid. The purpose was to evaluate the texts in terms of lexical and syntactic complexity as well as cohesion which included narrativity, syntactic simplicity, word concreteness, referential cohesion and deep cohesion. Each factor was able to increase or decrease the load of cohesion in the text. in addition, the output analysis managed to generally determine how difficult the passages were and what grade levels they suited. The findings revealed that there was some mismatch between the standard CEFR and NGSL wordlists and the number of words the learners study during their formal school language learning. Besides, the texts of the workbooks were to varying degrees incompatible with each other. Detailed examination uncovered some more features of each single text. Assessing the Iranian textbooks may be generalized to assist teachers, material developers, and writers determine what might present difficulties to language learners and then take actions to improve their reading and writing skills.

---

* Mohammad Abdollahi-Guilani, I got my PhD from the National University of Malaysia and I am interested in the subject of verb collocation in English and Persian. So far, I have published about 15 books and articles.

Readability Index and Reading Complexity in high school EFL Textbooks

Mohammad Abdollahi-Guilani *
(corresponding author)
Assistant Professor of English at Qazvin International University, Buin Zahra Center for Higher Technical and Engineering Education, Qazvin, Iran
Email: abdollahi20@gmail.com & abdollahi@bzte.ac.ir

ABSTRACT

This study aims to investigate the readability of Iranian English textbooks, Vision series 1-3, for high school. It focused on the reading sections of student books and workbooks. The texts were collected and fed into different readability tools such as Coh-metrix, and Flasch-Kincaid. The purpose was to evaluate the texts in terms of lexical and syntactic complexity as well as cohesion which included narrativity, syntactic simplicity, word concreteness, referential cohesion and deep cohesion. Each factor was able to increase or decrease the load of cohesion in the text. in addition, the output analysis managed to generally determine how difficult the passages were and what grade levels they suited. The findings revealed that there was some mismatch between the standard CEFR and NGSL wordlists and the number of words the learners study during their formal school language learning. Besides, the texts of the workbooks were to varying degrees incompatible with each other. Detailed examination uncovered some more features of each single text. Assessing the Iranian textbooks may be generalized to assist teachers, material developers, and writers determine what might present difficulties to language learners and then take actions to improve their reading and writing skills.

---

* Mohammad Abdollahi-Guilani , I got my PhD from the National University of Malaysia and I am interested in the subject of verb conjugations in English and Persian. So far, I have published about 15 books and articles.

38

A textbook can effectively support language learning and help language learners improve their skills and in this way the book also goes for careful evaluation in authentic teaching situations (Hakim, Setyaningsih, & Cahyaningrum, 2021). Therefore, a suitable textbook can channel and facilitate the process of education.

As Tomlinson (2008) mentions, a textbook serves as a point of reference for teachers to what extent their teaching has succeeded. In addition, textbooks can help the learners become familiar with the new things that will follow shortly (Wong, 2011). Likewise, according to Tomlinson (1998), the success of a textbook can be detected by at least 16 criteria three of which include their giving the learners a feeling of comfort, confidence, and mastery over the learning material.

Formal education is classically bound to a book and the reading content should be properly prepared to target the right reader in terms of difficulty. As Dubay (2004) noted, in evaluating a textbook, readability determines the difficulty of passages. According to Richards, *et al.* (1992, p. 306 as cited in Heydari & Riazi, 2012), readability means: "how easily written materials can be read and understood. There are several factors involved in readability like the grammatical complexity of the language in the text, the number of new words and the average length of sentences."

The readability of textbooks has widely been studied in different educational settings (e.g. Pitler & Nenkova, 2008; Crossley, Allen & McNamara, 2011), but to the best knowledge of present researcher, there is no research on readability of the reading section of the newly compiled *Vision* series in Iran. Therefore, this study may be the first tentative attempt to analyze these books especially because since their compilation, these books have been reportedly had their own strengths and weaknesses. However, the focus of this search is the quality of the reading sections in terms of readability and three readability tools are used: Flesch Reading Ease formula and Flesch-Kincaid Grade Level (Flesch, 1948) and more importantly, the Coh-Metrix L2 Reading Index (Crossley, Allen & McNamara, 2011) is used for the evaluation of cohesion of the texts. In addition, the wordlist of the Iranian English textbooks for secondary and high schools school

was used as a corpus for comparison with those of CEFR and NGSL.

To provide teachers and material developers with a better view of the quality of the reading passages of *Vision* series, this study aims to measure readability of this section by answering the following research questions

1. What is the level of vocabulary presented in Iranian English textbooks *Vision*?

2. What is the readability of the passages of the Iranian English textbooks *Vision*?

3. Is there a relationship between the passages of the student books and workbooks in terms of difficulty in the Iranian English textbooks *Vision*?

4. How cohesive are the reading passages of the Iranian English textbooks *Vision*?

2. Literature Review

For a long time, there had been no effort to edit or improve the quality of the textbooks both in secondary and high school levels. However, in 2012, after about a gap of two decades, a new series of books called *Vision* (Alavi-Moghaddam, et al. (2016) was introduced for high school. This series includes three student books and three accompanying workbooks for three years of high school. *Vision* book 1 contains four lessons and *Vision* books 2 and 3 have three lessons each.

Since their publication, these books have been variously reviewed from different aspects. Fazelimania, Donyaie, Yousefi, and Hafez (2019) evaluated *Vision One* through questionnaires and interviews, finding the respondents were generally satisfied with the book, but most of them found the book unsuccessful in fulfilling the set objectives.

Khodabandeh and Mombini (2018) collected their data through a questionnaire which targeted teachers and students in Khuzestan Province, Iran, to evaluate *Vision One* in terms of seven criteria namely, cultural considerations, language type, content, practical considerations, skills, and layout and design. They revealed that the subjects liked the book, but it was poor in integration of the cultural elements.

Similarly, Soodmand Afshar and Sohrabi (2020) examined *Vision Three* from different points of view but one of which was how cultural elements were integrated in the content of the book. They found that about 73.3 percent of the respondents believed that the target language

was not properly addressed in the book and localization of domestic culture had gone to the extreme as emphasized by 74% of the study teachers.

Similarly, Gheitasi, Aliakbari, and Yousofi (2020) examined the cultural load in the books by the culture categorization framework (Adaskou et. al., 1990) and cultural reference framework (Cortazzi & Jin, 1999), and found that the books mostly contained cultural content related to local culture with particular attention to the esthetic sense of culture as well as sociological, sociolinguistic and semantic senses.

Using Litz's checklist, Janfeshan (2018) collected the English teachers' views of Kermanhsha Province, Iran, about *Vision One*. Results indicated the teachers believed that the amount of exercises was not compatible with their limited time, which was aggravated by a high number of students in each class and lack of audio-visual aids.

Working on *Vision Three*, Saidi and Mokhtarpour (2020) asked 80 teachers to assess the book in terms of nine major criteria such as the four language skills as well as vocabulary, grammar, pronunciation, tasks and activities and supplementary materials. The findings revealed the teachers' satisfaction with the speaking and listening sections but they ranked the supplementary materials the least satisfactory.

And finally, the vocabulary the reading passages of *Vision* were also scrutinized in terms of patterns. Askani and Khatin Zadeh (2016) worked on the patterns of encoding motion events in the sentences and found the sentences were the simplified versions of typical English sentences and so did not offer a complete picture of such events. In addition, as for the use of metaphors to describe non-motion events in terms of motion events, the books underused metaphors.

Regarding the application of readability tools, Nahrkhalaji (2012) states that textbooks can be evaluated from various aspects such as credibility, validity, flexibility, authenticity, and practicality. In addition, there are some checklists and evaluation models that focus on the evaluation of textbooks in such angles as aims and appropriateness, graphics and flexibility, or affective processes and cognitive skills.

For the evaluation of texts, there are about 200 calculating tools and all of which primarily measure two aspects of the sentence: word difficulty and sentence difficulty (DuBay, 2004). The former is an analysis through word length or syllable count, and the latter is via the number of words in a sentence. And Wissing, Blignaut, and Van den Berg (2016) conclude readability formulas can measure appropriateness of a textbook.

Gyasi and Slippe (2019) examined the readability of three English textbooks for diploma students and suggested that all of the textbooks were generally in ranges between 'fairly difficult' and 'difficult' to read, which could give rise to the low readability of the books.

Owu-Ewie (2015) identified the readability of 48 reading texts of Junior High School and employed the Gunning FOG Readability test, the Flesch Reading Ease Formula, the Flesch-Kincaid Grade Level, Automatic Readability Index, and SMOG Index.

Coh-Metrix Tool was another tool that was used by Crossley, Greenfield, and Mcnamara (2008), who investigated cohesion and text problems at different stages of language discourse and theoretical evaluation.

There are very few studies on the readability of textbooks in the Iranian context. One was by Heydari and Riazi (2012), who focused on differences between the evaluation of EFL expert readers and computer-based evaluation of English text difficulty. Using a Likert-type scale to record their perception of the different components of text difficulty, the participants revealed significant differences between their evaluation of text difficulty and the Flesch Readability index of the texts.

Likewise, Mohebbi et al. (2017) investigated some ESP textbooks specialized for Iranian majors and found that most texts ranged from difficult to very difficult and hence unsuitable for post-secondary students in Iran.

As mentioned in the introduction, the Flesch Reading Ease formula and Flesch-Kincaid Grade Level (Flesch, 1948) are commonly used in readability evaluations. The Flesch Reading Ease Score shows a typical reader's approximate educational level to easily read and comprehend a particular text. The scores around 100 mean shows the text is extremely easy to read and around 0 mean represents a highly complex text and difficult to understand. For example, e.g., if the score is around 70 to 80 it means the text is appropriate for around school grade level 7, i.e., the average adult should find it reasonably simple to read.

40

The Flesch-Kincaid Grade Level also shows what educational level (in U.S.) a person will need to understand a particular text. For instance, if the text has a Flesch-Kincaid Grade Level of 9, the reader can understand the document easily and has completed around nine years of education.

Regarding a foreign language, there should be drawn a line between L1 and L2/ EFL readers. Grabe and Stoller (2012) point out these readers do not have the same language resources; their levels of lexical, grammatical, and discourse knowledge differ when they start reading; their exposure to native language print not only in terms of time, but also size is not equal and last but not least, L2/EFL readers have different social and cultural motivations for reading.

One problem with the Flesch-Kincaid readability scales is that they were originally devised for native speakers of English. Comparing other readability tools and CEFR (Common European Framework of Reference for Languages), which is a guideline for assessing the language ability of learners, Uchida and Negishi (2018) claimed most of the readability tools focused on English as a first language rather than a second or foreign one. They argued that the Flesch-Kincaid Grade Level basically indicates the complexity of the text based on sentence length and word complexity but does not consider word levels. For example, the words "cat" and "paw" that have one syllable each, but fall different difficulty levels of A1 in B2, respectively. Uchida and Negishi (2018), using the Japanese coursebook corpus, created an online system called CVLA (CEFR-based Vocabulary Analyzer) that assigns CEFR-J (i.e. Japanese version of CEFR) levels to a text to decide at what level the Japanese English textbooks stand according to CEFR classifications.

Other criteria such as the CEFR scales analyze documents for their accessibility to learners of EFL / ESL. CEFR classifies readers into 6 levels: A1, A2, B1, B2, C1, and C2. The highest or mastery level is C2. There are 15584 words and phrases and idioms in this classification and in each level there are some words that target a typical learner at a certain level.

Furthermore, the Flesch-Kincaid readability tool underestimates the difficulty for non-native speakers of documents containing short sentences. Besides, the scores cannot take into account many of the Latin-based complicated words in English that are not complicated for speakers of other Romance languages such as Spanish, Portuguese, French, and Italian. The fact is that the concept of difficulty is only partially measured from language to language and they may sound unfamiliar and incomprehensible for eastern language speakers such as Persian or Chinese natives. Table (1) gives a rough equivalent list of scores for different levels in Flesch-Kincaid measurement. In spite of such shortcomings, it is the most commonly used readability tool and can also be supported by other tools for specific purposes.

Table 1

Measurement Scores in Flesch-Kincaid Measurement and CEFR

Flesch-Kincaid readability scores and Grade Levels	CEFR
0-50 Very difficult (Higher education level)	C2 Mastery / Proficiency 8-9
50-60 Fairly difficult (11th or 12th grade-final years of high school)	C1 Advanced 7-8
60-70 Plain English-should be easy to understand by students from 14 to 15 years upwards	B2 Upper intermediate-6-7
70-80 Fairly easy-accessible to students aged 13 upwards	B1 intermediate-4.5-6
80-90 Easy	A2 elementary-3-4
90-100 Very easy :	A1 EFL beginners-1-2

Source: https://linguapress.com/teachers/flesch-kincaid.htm

3. Method

3.1. Data collection

The corpus of this study included two parts: words and reading passages. The words were derived from *Vision series* English textbooks for Iranian high school and *Prospect* Alavi-

Moghaddam, et al. (2013) for secondary school and the reading passages were collected from the *Vision* series including three student books and three accompanying workbooks. *Vision One* contained four lessons and *Vision Two* and *Vision Three* had three lessons each.

3.2. Procedure

A) The reading passages were scanned and prepared to be fed into online readability tools. Then, all the words used in the six books of high school books as well as the six books of secondary school were collected. The purpose was to have a basic corpus of all the lexicon that students studied during their formal school English learning. After that, all the words of 12 books were alphabetically listed. Proper names of people and places were excluded from the list, but nationality adjectives were kept unchanged. In addition, all regular plural or past and past participle forms were removed from the list but irregular forms (e.g. wife and wives; speak, spoke, spoken) were included to make a total of 3300 words as the final wordlist for Iranian English textbooks.

B) The words of the CEFR wordlist and NGSL were compared with the words the high school students faced in their school language learning to check whether the textbooks provided sufficient, useful and appropriate vocabulary items.

C) The texts were evaluated in terms of readability by means of some measurement tools like Flesch-Kincaid calculator.

D) The texts were also checked for cohesion by means of the Coh-Metrix Common Core Text Ease and Readability Assessor, which analyzes the "easability" and readability of texts in terms of five criteria as explained below:

1. Narrativity: By narrativity, it means to what degree the text is story-like. The more story-like a text is, the higher the narrativity score, and the easier the text.

2. Syntactic simplicity: Syntactic simplicity is achieved when the sentences have similar structures and verb tenses are easier to read in addition to fewer clauses, fewer words per sentence, and fewer words before the main verb.

3. Word concreteness: When there are more concrete words in a text, the text will be easier to read.

4. Referential cohesion: Referential cohesion takes place when there is an overlap between words, word stems, or concepts from one sentence to another. A text will be easier to read as long as the sentences and paragraphs have similar words or conceptual ideas.

5. Deep cohesion: Deep cohesion occurs when there are enough transitional words to link events, ideas and information of the whole text together.

After the analysis, the research questions were answered.

4. Results and Discussion

4.1. Results

First of all, the list of words that Iranian EFL learners studied in their books was compared with that of CEFR list assigned for EFL leaners. In CEFR wordlist, there are 15271 words and phrases. Some of the words may be duplicates but their usage / function is different. As mentioned earlier, this list comprises six levels of English vocabulary knowledge of EFL leaners as seen in Table (2).

Table 2

CEFR Wordlist and Iranian English Textbook Wordlist for School

CEFR Levels	# of words in CEFR	# of words in *Prospect* and *Vision*
A1	666	533
A2	1564	784
B1	2889	1052
B2	4128	897
C1	2349	340
C2	3675	342
Total	15271	3948

42

As Table (2) shows, there are 15271 words in CEFR wordlist, but there are only 3948 words (in fact, entries) contained in the whole 12 books used in the six years of secondary and high school levels in Iran. These figures initially represent lack of equivalence between the standard wordlist and Iranian school English lexicon. A quick look at the figures indicates that for example, as for CEFR A1 Level, which consists of 666 words, there are 533 words in Iranian wordlist that fall into this category. But as the level of CEFR goes higher, there is a decrease in the number of words that Iranian EFL learners require in order to meet the international standard of vocabulary knowledge. When it goes to C2 Level, where over 3600 words are needed, Iranian students study less one sixth of those words. This imbalance can signal that they Iranian EFL leaners are inadequately equipped with vocabulary knowledge which will directly affect reading comprehension.

This study aimed to study the readability of Iranian high school English textbooks with regard to the reading passages given both in the student books and workbooks in three levels. Table (3) shows lexical density of the passages in the *Vision* series in student books and workbooks. By lexical density, it is meant the ratio of lexical words (or content words including nouns, adjectives, verbs, and adverbs) to the total number of words used in the text (Bayazidi, Ansarin, & Mohammadnia, 2019). When there is a high percentage of lexical density in a text, as Didau (2013) believes, that text is more informative. For example, in the sentence "The quick young runner suddenly jumped over the wide river.", there are 7 lexical words out of 10, meaning that lexical density of the above sentence is 70%. A lexically dense text typically scores at around 56% or above and the message is clear, but in the sentence "He said to her he helped her.", because of the overuse of pronouns (or function words), there is a high chance of misinterpretation of the message with a lexical density of 28.57%.

Table 3

Analysis of Word Class per Passage for Vision Series (Based on New-GSL)*

Lessons	Vision 1				Vision 2			Vision 3		
	1	2	3	4	1	2	3	1	2	3
size(s)	155	168	213	192	296	312	310	305	290	343
size(w)	162	170	144	150	252	278	264	178	274	272
N(s)	29/45	33/44	39/42	40/51	48/68	77/86	69/93	47/60	48/69	92/104
N(w)	33/37	13/27	39/39	27/32	49/60	62/77	51/55	18/22	48/70	47/69
V(s)	25/25	26/28	43/45	21/22	33/36	61/63	39/40	52/61	53/53	60/67
V(w)	31/31	33/35	29/32	23/25	53/56	49/54	46/49	33/38	42/46	45/51
Adj(s)	4/8	21/21	14/15	17/22	13/21	31/35	22/35	14/16	11/15	18/37
Adj(w)	5/5	17/19	10/11	13/14	16/18	25/32	26/28	9/9	12/19	13/32
Adv(s)	8/9	4/4	18/21	2/3	15/17	16/20	8/10	21/28	18/19	11/13
Adv(w)	10/15	14/15	6/7	19/19	7/8	11/12	14/14	16/17	13/13	11/11
Con(s)	23/23	25/25	33/33	34/34	56/56	33/33	65/65	31/31	37/37	56/56
Con(w)	20/20	26/26	22/22	15/15	35/35	40/40	44/44	24/24	45/45	33/33

*S: Student book, W: Workbook, N: Noun, V: Verb, Adj: Adjective, Adv: Adverb, Con: Connectors and prepositions Connectors

The New General Service List (NGSL) refers to the core high frequency vocabulary words for students of English as a second or foreign language. That is a major update of Michael West's 1953 General Service List which was limited in scope and was partly out dated. This new list is based on a carefully selected 273 million-word subsection of the 2 billion word Cambridge English Corpus and it includes words that are frequently used by native speakers of English in general contexts (Brezina, & Gablasova, 2015).

As Table (3) shows, there is a detailed analysis of the word classes used in *Vision* student books and workbooks compared with the New-GSL wordlist. The first two rows show how many words comprised each passage. A quick look indicates that size of the reading passages in the student books and their corresponding

counterparts in the workbooks are almost the same. However, the words in all lessons of the student books except for Lessons One and Two outnumber those of the workbooks.

The word classes that were compared included only five parts of speech as they are mostly employed in texts. For each word class, two numbers were recorded separated by a forward slash. For example, in the first column for *Vision One*, Lesson One, there is a number 29/45 for the cell for nouns. This figure means that in the general list there are 29 nouns, but the text had 45 nouns. That is to say, 16 words were off the list. Browsing through the results suggests that other parts of speech used in the passages of the books have almost a close number to the *NGSL* figures. When compared with the figures given for the nouns in the workbook, it appears to have more words off the list, possibly representing more difficult texts in the student book passages. This feature is not true for other parts of speech. It may resonate the idea that more nouns are needed in conveying messages than are other word classes. On the contrary, a simple comparison can be made on the connectors used. They are the same. Perhaps, this is because such words belong to a closed class of function words while nouns are content or lexical words and give a text its meaning and provide information regarding what the text is about.

It is noteworthy that in NGSL there are 2230 words (164 duplicates have been removed). In the Iranian English lexicon for school there are roughly 3262 words. A comparison was made between these two lists and the result showed that the core words from words in NGSL that were used in the Iranian corpus comprised 1460 words. It means that about 65% of the words that were quite frequent were provided to the learners. In other words, although the words in the Iranian corpus outnumbered those of NSGL, Iranian EFL learners still lacked about 35% of the words that were essential in language learning.

Another measuring tool used for the readability of the *Vision* series was Flesch Kincaid Calculator. This tool counted the number of words and sentences in the passage, number of words in each sentence, and average syllables per word. Then based on this information it yielded a score which showed how easy that text was, what grade level it had and what age group could read it. Table (4) is the result of Flesch Kincaid analysis for the reading passages of *Vision* student books and workbooks.

Table 4

Readability of Vision Series

Vision and lessons	# of words	# of sentences	Average words per sentence	Average Syllables per Word	Flesch Reading Ease Score	Flesch-Kincaid Grade Level	Reading Level
S:V1,L1	158	15	10.5	1.4	77.7	5	7th
W:V1,L1	162	12	13.5	1.5	66.2	7.4	8th-9th
S:V1,L2	168	18	9.3	1.3	87.4	3.4	6th
W:V1,L2	170	18	9.5	1.6	61.8	7	8th-9th
S:V1,L3	213	19	11.2	1.4	77	5.3	7th
W:V1,L3	144	16	9	1.5	70.8	5.6	7th
S:V1,L4	192	18	10.7	1.7	52.2	8.6	10th-12th
W:V1,L4	150	11	13.6	1.5	66.1	7.4	8th-9th
S:V2,L1	296	22	13.5	1.6	57.8	8.6	10th-12th
W:V2,L1	252	25	10.1	1.5	69.7	6	8th-9th
S:V2,L2	312	25	12.5	1.6	58.8	8.2	10th-12th
W:V2,L2	278	17	16.4	1.6	54.8	9.7	10th-12th
S:V2,L3	310	19	16.3	1.6	54.9	9.6	10th-12th
W:V2,L3	264	15	17.6	1.5	62.1	9	8th-9th
S:V3,L1	305	22	13.9	12	82.7	5.2	6th

44

Vision and lessons	# of words	# of sentences	Average words per sentence	Average Syllables per Word	Flesch Reading Ease Score	Flesch-Kincaid Grade Level	Reading Level
W:V3,L1	178	10	17.8	1.3	78.8	6.7	7th
S:V3L2	290	24	12.1	1.7	50.7	9.2	10th-12th
W:V3L2	274	18	15.2	1.7	47.6	10.4	College
S:V3L3	343	26	13.2	1.6	58.1	8.4	10th-12th
W:V3L3	272	18	15.1	1.5	64.6	8	8th-9th

As Table (4) displays, in *Vision One*, there is a partial balance between the number of sentences in the passages of the student books and workbooks (i.e. 15-12, 18-18, 19-16 etc.), but when it comes to *Vision Three*, the sentences in the student book outnumber those in the workbooks as do the total words in each passage. As for other features, such as average words per sentence and average syllables per word, the numbers are close to each other. However, in terms of easibility there does not seem to be a steady trend. For example, reading ease score for student book S:V1,L1 is 77.7 which can be assigned to Reading Level 7, while its counterpart W:V1,L1 has 66.2 and stands in the 8th-9th reading level. This is true for most of the other lessons, indicating that the workbook looked more difficult, but for S:V1,L4 and W:V1,L4 and S:V2,L1 and W:V2,L1, S:V3L3 and W:V3L3 the results are reverse. and the reading passages in the workbooks are easier than those in the student books, although there is a sheer incongruity in the reading levels of V3,L1 with the scores of 6 and 7 compared with the reading levels of other lessons of other books which have higher scores before and after this level.

A major issue in the measurement is that when the result of the Reading Level assigns a number, it means that the reader has already studied English for that number of years. For example, in Table (4) the last cell in the first row, the reading level is 7 and it means that the student has already studied English for 7 years, but Iranian students are formally exposed to English only for three years before high school. This may suggest that the English material that Iranian school students intensively learn is almost double that of a native learner in seven years. With this knowledge in mind, it may be claimed that Iranian EFL learners are doing 12 years of native speakers' learning of English in a period of 6 years.

Another factor that was investigated was cohesion. It was analyzed by means of Coh-Metrix L2 Tool. Table (5) presents some information about the "easability" and readability of the reading passages both in the student books and workbooks of *Vision* Series, Iranian English textbooks for high school. The classification is in terms of five criteria namely narrativity, syntactic simplicity, word concreteness, referential cohesion, and deep cohesion. The analysis shows what component of the text makes it cohesive or less cohesive.

Table 5

Cohesion of the Vision Reading Passages by Coh-Metrix

Sources	Narrativity	Syntactic Simplicity	Word Concreteness	Referential Cohesion	Deep Cohesion	Flesch-Kincaid Grade Level
Student: V1L1	45	81	87	9	47	6
Work: V1L1	57	68	97	63	15	7
Student: V1L2	33	66	73	63	2	4
Work: V1L2	60	74	4	83	45	5
Student: V1L3	81	75	42	28	67	6
Work: V1L3	53	93	93	37	79	6

Student: V1L4	5	78	72	41	35	8
Work: V1L4	80	50	39	67	94	7
Student: V2L1	55	37	11	76	8	8
Work: V2L1	50	88	72	31	78	5
Student: V2L2	32	66	64	17	50	7
Work: V2L2	36	46	72	21	54	10
Student: V2L3	20	37	92	75	63	9
Work: V2L3	55	23	5	80	50	8
Student: V3L1	93	57	98	81	82	5
Work: V3L1	100	29	13	95	100	7
Student: V3L2	38	93	4	56	54	8
Work: V3L2	32	78	11	51	58	8
Student: V3L3	14	90	81	50	82	8
Work: V3L3	22	55	97	72	24	8

As Table (5) shows, the first Lesson of both books (i.e. student book and workbook) of *Vision One* has the following features. Both are average in narrativity. The student book is high in syntactic simplicity, meaning that it has simple sentence structures and so the text is easier to process, while the workbook is average in this regard. Both of them enjoy high word concreteness, suggesting a low volume of word abstractness, hence high imageability and easier understandability. The student book has low referential cohesion. Thus, there is less overlap in explicit words and ideas between sentences. These conceptual gaps require the reader to make more inferences, but the amount of referential cohesion is average for the workbook. The amount of deep cohesion for the student book is average, while this is low for the workbook, suggesting a lack of explicit causal relationships when needed by the text. Because of this, it may be more difficult to comprehend on unfamiliar topics. All of these lead to more difficulty of Lesson One of the workbook compared to that of the student book with a score of Flesch-Kincaid Grade Level 7 against 6.

As for Lesson Two, both texts are average in narrativity and syntactic simplicity. The student book is average on word concreteness but this feature is low for the workbook suggesting a high volume of word abstractness and low imageability. Thus, it may be more difficult to understand. The amount of referential cohesion for the student book is average while this is high for the workbook, displaying more overlap in explicit words and ideas between sentences. The lack of conceptual gaps requires the reader to make fewer inferences. For the student book, deep cohesion is low, suggesting lack of explicit causal relationships when needed by the text. And this may make the text more difficult to comprehend. This component is average for the workbook. The Grade Levels for the student book and workbook are 4 and 5, respectively.

Lesson Three: When narrativity is high (as is for the student book), it means that it is more story-like. More story-like texts are usually easier to comprehend. In terms of syntactic simplicity, the student book is average, but the workbook is high, meaning that it has simple sentence structures; hence, easier to process. Word concreteness is average in the former but high in the latter book. Both enjoy an average amount of referential cohesion. Regarding deep cohesion, however, the student book is average and the workbook is high, suggesting more explicit causal relationships as needed by the text. Because of this, it may be easier to comprehend on unfamiliar topics. In total, different features supersede in different aspects and yield an identical Grade Level of 6.

Lesson Four in each book shows opposite stances in narrativity: low and high. Low narrativity signals less comprehensibility of the text, but it may be counterbalanced by its high syntactic simplicity. Both texts have average word concreteness and referential cohesion. But deep cohesion is high for the workbook, indicating that it may be easier to comprehend on unfamiliar topics. All in all, the student book stands on Grade Level 8, while the student book

46

on 7. For *Vision One*, it can be concluded that different elements stick together and help the reader comprehend a text. As shown, the passages are almost at the same level although in some features they may pose challenges for the readers. Their grade levels are also similar.

Vision Two:

Lesson One: Although both books enjoy average narrativity, there is a higher level of syntactic simplicity in the workbook, which makes the text easier to process. Word concreteness is low in the student book and this suggests a high volume of word abstractness and low imageability. In addition, a high level of referential cohesion in the student book makes the text easier to comprehend. However, in terms of deep cohesion, the student book is low and this suggests lack of explicit causal relationships when needed by the text. Because of this, it may be more difficult to comprehend on unfamiliar topics. The wide gap between different components in the two books especially in terms of deep cohesion gives rise to a wider gap in Grade Levels: Student book (8) vs workbook (5), making the student book much more difficult to read. This can sound like a sudden impact on the learners when they enter a new academic year with a new book.

Lesson Two in both books has average narrativity and syntactic simplicity. About word concreteness in the student book, it is low but average in the workbook. As for referential cohesion, there is a higher level in the student book and so the reader has to make fewer inferences in the student book. Regarding deep cohesion, the reader encounters a low level in the student book, suggesting lack of explicit causal relationships when needed by the text. Because of this, it may be more difficult to comprehend on unfamiliar topics. Like the status of the grade levels in Lesson One between the student book and the workbook, in Lesson Two the condition is the same but quite reverse in position. That is to say, the grade level of the workbook (#10) outranks the student book (#7). These mismatches may cause some challenges, if not called obstacles, when the learners are doing the workbook by themselves as is common in the Iranian setting.

The only common features that these two books have in Lesson Three are the referential cohesion and deep cohesion which are high and average, respectively. Due to degree of some components such as word concreteness in the workbook, which is about 5%, the text becomes more difficult to understand. There is not a significant difference in grade levels between these books in this Lesson.

Vision Three

As Table (5) shows, in Lesson One of *Vision Three*, narrativity is high and syntactic simplicity is average in both books and this can represent an easily comprehensible passage. In addition, referential cohesion and deep cohesion are high, too. It means that there is more overlap in explicit words and ideas between sentences. The lack of conceptual gaps requires the reader to make fewer inferences. Besides, it may be easier to comprehend on unfamiliar topics. All these four components can show that the text is easy to understand. However, word concreteness is high in the student book, suggesting a low volume of word abstractness and high imageability and this is opposite to those in the workbook; hence, more challenges for the reader. As a result, the Grade Levels are 5 and 7 for the student book and workbook, respectively.

Lesson Two of *Vision Three* may be the best match among all of these books. These texts (i.e. student book and workbook) are average in narrativity. Their high syntactic simplicity means that they have simple sentence structures, which means easier processing. They have low word concreteness, suggesting a high volume of word abstractness and low imageability. Thus, they may be more difficult to understand. They have an average amount of referential and deep cohesion. Both texts lie in Grade Level 8.

Although Lesson Three in both books fall in the same Grade Level of 8, there are minor differences between them in syntactic simplicity (high in the student book and average in the workbook). In addition, a high deep cohesion level can be observed in the student book and a low one in the workbook. Above all, these two books have an overall identical grade level.

4.2. Discussion

To answer the first research question, What is the level of vocabulary presented in Iranian English textbooks, *Vision*?, following the data given in Table (2) about CEFR wordlist, Iranian English textbooks cannot provide sufficient vocabulary for the leaners to help them confidently accompany the standard level of reading comprehension which is mostly based on word knowledge. On the contrary, compared with the New General Service wordlist as given in Table (2) above, it can be concluded Iranian EFL learners are bombarded with more words than the list suggested, but firstly, they are less frequent and secondly they do not comprise at

least 35% of the words classified as core words with high frequency.

The second research question asked about the readability of the passages of the Iranian English textbooks, *Vision*. To answer this question, different tools were used and most important of all were Flesch-Kincaid tools. With respect to the fact that the tools had originally been designed for native learners of English and not second / foreign leaners and considering that Iranian learners cram all the 12 years of knowledge of English learning in 6 years, and finally, given the fact that native speakers are frequently exposed to language both orally and visually, the textbooks had varying degrees of readability, but with rough scores. The findings show that there was no consistency in arrangement of reading passages. In some cases, the workbooks dominated the student books in terms of size of the reading passages, number of sentences, words in the sentences and syllables in the words. And sometimes, the dominance was the way round. Therefore, there cannot come out a definite decision to the readability of the passages.

The above explanation can also answer research question three which asked if there was a relationship between the passages of the student books and those of workbooks in terms of difficulty in the Iranian English textbooks, *Vision*. A simple reply is a 50-50 relationship, because in some aspects there are some features that outweigh the other as explained in answer to the fourth research question below.

Based on the findings, and considering the fact that English is a foreign language in Iran, it could roughly be said that two lessons in the student book were easy; one was Lesson 2 in *Vision* 1 and the other Lesson 1 in *Vision* 3. Four lessons were fairly easy: Lessons 1 of the student book in *Vision* 1 and lesson 3 of the student book and workbook in *Vision* 1, and Lesson 1 of the workbook in *Vision* 3. The standard level included 6 lessons from the workbooks: Lessons 1, 2 and 4 in *Vision* 1, Lessons 1 and 3 in *Vision* 2 and Lesson 3 in *Vision* 3. In total, seven lessons were ranked as fairly difficult five of which included the reading passages in the student book: Lesson 4 in *Vision* 1, Lessons 1 of the student book and lesson 2 (plus the workbook passage) and lesson 3 in *Vision* 2 and lessons 2 and 3 in *Vision* 3. And finally Lesson 2 of the workbook in *Vision* 3 was difficult; in fact, it was at the college level. It seems that the distribution of texts in

terms of difficulty was not hierarchically considered throughout the 6 books.

Findings of the current study indicated that out of 20 passages studied in *Vision* series of Iranian EFL textbooks for high school, only two texts were easy, four fairly easy, seven fairly difficult, six at the standard level and one was difficult. These anomalies seem to be in line with the results of many studies. Owu-Ewie (2015) studied English textbook readability in Ghana and learned that many of the texts were inappropriate with the level of the students, which was above their level. Another case (Gyasi & Slippe, 2019) revealed that three textbooks being analyzed were generally between 'fairly difficult' and 'difficult' levels, giving rise to the low readability of the textbooks for diploma students. Moreover, Tasaufy's (2017) study suggested that only three texts out of nine suited the level of the students. In the same vein, Miftaahurrahmi and Syarif. (2017) on ten samples of texts and noticed that only one text actually suited the grade level of their students.

But the final question was: How cohesive are the reading passages of the Iranian English textbooks, *Vision*?

To answer this question, the analysis of the passages by Coh-Metrix could give a more definite response. Based on the results, it can be concluded that narrativity of the passages in student books started with 45% in Lesson One of *Vision One* declined in Lesson Two, then rose to 81% in Lesson Three. Suddenly in Lesson 4 there was a sharp fall to 5%. As for the workbook for this level, there was a steady trend in the level of narrativity starting with 50% on average but only in Lesson Three, it went up to 80%. It can be concluded that there was more consistency in narrativity in the workbook.

The syntactic simplicity of the reading passages (in student books) started with a percentage of 81 and fell down and then rose to 78%. But this trend was almost reverse in the workbook where it began with 68% and rose to 93 and in Lesson Four it declined to 50%. Regarding *Vision* Two, syntactic simplicity was at a low level (with 37%) then increased to 66% and again decreased to 37%. With regard to the workbook, the first lesson had a high degree of syntactic simplicity, but for the following lesson there was a sharp decline reaching at 23% in Lesson Three. *Vision Three* which should logically be a harder level with less syntactic simplicity, started with 57%

but peaked to 93%. It can be concluded that there was less syntactic simplicity in the workbook.

As Ortega (2003) found out, a text can be less syntactically simple if there is a high degree of sophistication (i.e. the frequency of certain words) in language production as this can happen by such subcomponents as production length, sentence complexity, subordination, coordination, and the use of particular grammatical structures (Bulté & Housen, 2012).

Based on the results of Coh-Metrix analysis on word concreteness, the lowest level of concreteness was observed in 6 lessons, which makes the reading of those texts difficult, but in general all the other lessons enjoy a high level of this feature with a percentage ranging from about 40 to 97.

Excluding the first and third lessons of *Vision One*, and underestimating the second lesson of *Vision Two*, there is a high level of referential cohesion in the books and this signifies an above average level of cohesion in these books.

Perhaps, deep cohesion can finally determine the overall cohesion of the text, making the text easier to read and comprehend. Except for two lessons with critically low scores, the whole series enjoys an above average percentage of deep cohesion.

And finally, the Flesch-Kincaid Grade Level formula yields an average score of 7 for both books: 6.9 for the student book and 7.1 for the workbook. And this may conclude that *Vision* English textbooks for Iranian for high school can satisfactorily meet the requirements of Iranian EFL leaners.

5. Conclusion

Although there are many readability formulas to measure the difficulty levels of texts, they may not be totally trusted as they may fail to address some seemingly minor features. One of them is the perception of the reader which cannot be objectively calculated by counting syllables, word length, sentence length, and other text characteristics. The readers' age, and their linguistic and metalinguistic / metacognitive prior knowledge are not taken into consideration. The readers' personal reading skills including comprehension, pace of reading, and reading and re-reading strategies are among the factors that may

play a role in the readability of a text. For example, the sentence "It was covered with snow, so they stayed indoors.", may sound either shocking or welcoming to different readers with different world knowledge: A skier prefers to go out, but an elderly person may prefer sitting by the fireplace.

Despite their many supporters, readability formulas have been under criticisms because only sentence length and word difficulty were taken into consideration and these are not totally accurate and useful criteria (Bailin & Grafstein, 2001). These formulas also fail to fully address such factors as text structure, complexity of ideas, and schemata as well as readers' motivation and purpose for reading (Zamanian & Heydari, 2012). Furthermore, they are unable to deal with content difficulty and familiarity, the organization of ideas, the authors' style, and page layouts (Armbruster *et al.*, 1985, as cited by Odo, 2018).

Although cohesion can, to a great extent, be evaluated by tools such as Coh-Metrix, there are other peripheral factors that play key roles in this assessment. And finally, Bailin and Grafstein (2001) believe there is a peril that teachers' professional judgements may be superseded by these seemingly objective statistics.

Different analyses presented in the research can help teachers, material developers, and writers look into a text more deeply in order to determine what might present difficulties for language learners and then they can take appropriate actions to pave the road for better mastery of reading and even writing skills. Further research may be needed on the readability of the conversations of the textbooks and their relationship with the readability of the reading passages. Furthermore, the relation between mastery of CEFR wordlist and the psychological overload of language learning can be another topic for further research.

When compared with the CEFR wordlist, it was noticed that the Iranian English textbook was extremely flooded with local proper and place names. One major role in learning a new language may be played by self-identification with the atmosphere of the language being learned. Thus, it is suggested that further focus be

placed on the comparative study of names in EFL books.

Although the current study made use of Coh-Metrix for the evaluation of reading passages, it can be professionally used to check whether students' writing products have strengths and weaknesses in use of textual features and whether they conform to writing strategies and styles.

Despite the significant results of this study, there have been some limitations. One of the issue was the size of each passage in the books which were on average 200 words long. If the texts are longer, there may be more reliable results to allow for the detection of the effects of syntactic and lexical complexity and syntactic features. In readability and the ensuing comprehensibility of English as foreign or a second language, the scaffolding effects of first language cannot be ignored. When English is the first language, or a language of Romance-based family, text proficiency is undoubtedly higher, which did not take place in this research.

Adaskou, Kheira, Donard Britten, and Badia Fahsi. (`1990) "Design decisions on the cultural content of a secondary English course for Mororcco." *ELT journal* 44, no. 1 (1990): 3-10.

Alavi-Moghaddam, S. B. et al. (2013). *Prospect*. Iran: Tehran. Also available online at www.chap.sch.ir.

Alavi-Moghaddam, S. B. et al. (2016). *Vision*. Iran: Tehran. Also available online at www.chap.sch.ir.

Askani, S., & Khatin Zadeh, O. (2016). Analysis of Lexical Patterning and Semantic Elements in Reading Comprehension Sections of EFL Textbooks. *Journal of Modern Research in English Language Studies*, *3*(4), 25-1.

Bailin, A., & Grafstein, A. (2001). The linguistic assumptions underlying readability formulae: A critique. *Language & Communication*, *21*(3), 285–301.

Bayazidi, A., Ansarin, A. A., & Mohammadnia, Z. (2019). The relationship between syntactic and lexical complexity in speech monologues of EFL learners. *Applied Research on English Language*, *8*(4), 473-488.

Brezina, V., & Gablasova, D. (2015). Is there a core general vocabulary? Introducing the new general service list. *Applied Linguistics*, *36*(1), 1-22.

Bulté, B., & Housen, A. (2012). Defining and operationalizing L2 complexity. In A. Housen, F. Kuiken & I. Vedder (Eds.), *Dimensions of L2 performance and proficiency: Complexity, accuracy and fluency in SLA* (pp. 23-46). Amsterdam/Philadelphia: John Benjamins.

Cortazzi, M., & Jin, L. (1999). *Cultural mirrors: Materials and methods in the EFL classroom* (pp. 196-219).

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.

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

Didau, D. (2013), Black space: improving writing by increasing lexical density, from *The Learning Spy: Brain Food for the Thinking Teacher.*

Dubay, W. H. (2004). *The Principles of Readability*. Online Submission.

Fazelimania, A., Donyaie, S., Yousefi, M., Ranjbar, N., Hafez, F. (2019). An Evaluation of *Vision 1: An English Textbook for Grade 10. Critical Studies in Texts & Programs of Human Sciences*, 19(6), 273-297. doi: 10.30465/crtls.2019.4387

Flesch, R. (1948). A new readability yardstick. *Journal of Applied Psychology, 32*(3), 221–233. https://doi.org/10.1037/h0057532.

from http://lrr.modares.ac.ir/article-14-16583-en.html. Accessed August 10, 2021

Gheitasi, M., Aliakbari, M., & Yousofi, N. (2020). Evaluation of culture representation in *Vision* English textbook series for Iranian secondary public education. *Evaluation*, *12*(26).

Grabe, W., & Stoller, F. L. (2012). Teaching reading: Foundations and practices. *The Encyclopedia of Applied Linguistics*, 1-9.

Gyasi, W. K., & Slippe, D. P. (2019). Readability of English language textbooks for diploma students of the University of Cape Coast. *International Journal of Research*, *8*(1), 107–115.

Hakim, A. A., Setyaningsih, E., & Cahyaningrum, D. (2021). Examining the readability level of reading texts in English textbook for Indonesian senior high school. *Journal of English Language Studies*, *6*(1), 18-35.

Heydari, P., & Riazi, A. M. (2012). Readability of texts: Human evaluation versus computer index. *Mediterranean Journal of Social Sciences*, *3*(1), 177-177.

Janfeshan K. (2018) Investigating Iranian new senior high school book (*Vision* 1) based on book evaluation checklist from Kermanshah EFL teachers' perspective. LRR. 2018; 9 (5): 197-239.

Khodabandeh, F., & Mombini, R. (2018). Iranian EFL teachers' and students' perceptions towards the first grade high school English textbook (*Vision*1). *Journal of English Language Pedagogy and Practice*, *11*(22), 141-167.

Miftaahurrahmi, F., & Syarif, H. (2017). The readability of reading texts in English textbooks used by senior high school students in West Sumatera. *Advances in Social Science, Education and Humanities Research (ASSEHR)*, 199-203.

Mohebbi, H., Nayernia, A., Nemati, M., & Mohebbi, B. (2017). Readability of ESP textbooks in Iran: A neglected issue or a taken-for-granted one? *Journal of Teaching English for Specific and Academic Purposes*, *4*(3), 641–654.

Nahrkhalaji, S. S. (2012). An evaluation of a global ELT textbook in Iran: A two-phase approach. *International Journal of Humanities and Social Science*, *2*(3), 184-191.

Odo, D. M. (2018). A comparison of readability and understandability in second language acquisition textbooks for pre-service EFL teachers. *Journal of Asia TEFL*, *15*(3), 750.

Ortega, L. (2003). Syntactic complexity measures and their relationship to L2 proficiency: A research synthesis of college-level L2 writing. *Applied Linguistics, 24*(4), 492-518.

Owu-Ewie, C., & Eshun, E. S. (2015). The use of English as medium of instruction at the upper basic level (Primary four to junior high school) in Ghana: From theory to practice. *Journal of Education and Practice*, *6*(3), 72-82.

Pitler, E., & Nenkova, A. (2008). Revisiting readability: A unified framework for predicting text quality. In Proceedings of the 2008 conference on empirical methods in natural language processing (pp. 186-195).

Saidi, M., & Mokhtarpour, H. (2020). Can *Vision* 3 notch up a win over pre-university English textbook?. *Language Teaching Research Quarterly*, *15*, 21-32.

Soodmand Afshar, H., & Sohrabi, A. (2020). A critical evaluation of *Vision 3*: Rights analysis and culture in the spotlight. *Journal of Language and Translation Studies (JLTS)*, *53*(4), 133-97.

Tasaufy, F. S. (2017). The readability level of the reading texts in english textbook entitled 'bahasa inggris x'. *Edulitics (Education, Literature, and Linguistics) Journal*, *2*(2), 62-69.

Tomlinson, B. (1998). *Materials Development in Language Teaching* (B. Tomlinson (ed.); Second Edi). Cambridge University Press.

Tomlinson, B. (2008). *English language learning materials: A critical review*. Bloomsbury Publishing.

Uchida, S., & Negishi, M. (2018). Assigning CEFR-J levels to English texts based on textual features. In *Proceedings of the 4th Asia Pacific Corpus Linguistics Conference* (pp. 463-468).

Wissing, G. J., Blignaut, A. S., & Van den Berg, K. (2016). Using readability, comprehensibility and lexical coverage to evaluate the suitability of an introductory accountancy textbook to its readership. *Stellenbosch Papers in Linguistics*, *46*, 155-179.

Wong, P. W. L. (2011). Textbook evaluation: A framework for evaluating the fitness of the Hong Kong New Secondary School (NSS) curriculum. In *MA dissertation, Hong Kong: City University of Hong Kong*.

Zamanian, M., & Heydari, P. (2012). Readability of Texts: State of the Art. *Theory & Practice in Language Studies*, *2*(1).