Developing a Critical Checklist for Textbook Evaluation

Siavash Zokaeieh[1*], Shaban Najafi Karimi[2], Farzaneh Nouri[3], Afsaneh Hakimzadeh[4]

1, 2, 3, 4. Department of English , Qaemshahr Branch, Islamic Azad University, Mazandaran, Iran
*Corresponding author: szokaieh@gmail.com

....

Abstract

This study has been carried out to develop a critical checklist for global/commercial textbooks which play a crucial role in language teaching/learning. For this aim, a number of items have been developed based on a comprehensive review of the related literature and experts' opinions. The tentative checklist was administered to the targeted population, yet 326 checklists were deemed appropriate for the analysis. The statistical analyses utilized in this study were exploratory factor analysis and partial confirmatory factor analysis. Prior to the exploratory factor analysis, a correlation analysis reduced the items into 44. Next, the exploratory factor analysis identified ten components. Additionally, six items were excluded from the checklist due to simultaneous factor loading and the finalized checklist was truncated into 38 items. To investigate the possibility for further validation studies, the result of partial confirmatory factor analysis was not statistically significant. Results of this study have the potentials to shed some lights on the ideological aspects of textbook evaluation.

Keywords: evaluation, checklist, critical, developing**,** textbook

Introduction

Educational materials such as textbooks, videos, and realia in language classrooms play an irresistible role today. Textbook development is a profitable industry for the economy of the dominant language countries around the world (see Hadley, 2014). Different political and cultural agendas can also disperse the world via global textbooks. On the other hand, teachers, learners, test developers, policy makers, curriculum designers rely on textbooks as the core property in many circumstances. Therefore, identification of weakness and strength of textbooks as a linchpin from two main vantage points namely learner and teacher is fundamentally important.

Regarding the historical consideration of learners' advantage, O'Neil (1982) mentioned that textbooks could possibly cause situations to direct courses along with the opportunity to catch up with missed classes and prepare language learners before attending a session. This view is suitable for the banking model of education in which good learners are defined by their capabilities over accomplishment of certain globally selected lessons and tasks. In accordance to course direction features of global textbooks, Davies (2006) believed that locally planned syllabuses were more useful. More considerate approaches to the role of materials were elaborated by Allwright (1981) through distinguishing between deficiency and difference views. In deficiency view, the role of materials is to support learners for possible shortage such as teachers' epistemic knowledge. On the other hand, difference view bolds the role of experts (theorists and theory-based approaches to material development) in contrast to classroom teachers in terms of teaching materials. Allwright (1981) emphasized the learner-education, which tried to engage and involve learners not only in the classroom but also in decision making. As a result, learners become empowered and this empowerment may to certain degree encompass the concerns for inclusion of global dictation. This splendid approach is in line with the characteristics of postmethod learners which consider decision making and learners' autonomy as the constructive elements of pedagogy

(see Kumaravadivelu, 2006). Additionally, Tomlinson (2012) mentioned that "we need textbooks to save time and money and many teachers want a coursebook which provides everything they need in one source" (p. 158). Time and money remain two salient and undeniable factors, but they should not hold the ground in expense of fainting intellectual aspects of teaching/learning. These two factors might manipulate good teachers as followers of the dictated methods in educational settings (O'Neil, 1982).

From teachers' vantage points, Garinger (2002) accentuated the helpfulness of the textbook in that novice teacher didn't need developing the materials on the limited time and budget. Also, textbooks are believed to play a crucial role when English teachers are incompetent to deliver materials properly (Edge & Wharton, 1998). Teacher's job becomes fairly straight forward as they follow sets of predetermined, graded and sequenced lesson plans and objectives. In contrast, Hutchinson and Torres (1994) stated that "the danger of ready-made textbooks is that they can seem to absolve teachers of responsibility" (p.315). This is in line with Ur's (1996) assumption that textbooks might faint the creativity and impetus of teachers by over-reliance on prefabricated materials.

However, in the perplexed views of proponents and opponents, we cannot undermine the fact that textbooks are firmly entrenched in education. Hadley (2014) stated that global textbooks "can play an important role in helping, and not harming, second language learning" (p.230). Therefore, investigation and evaluation of textbooks from diverse points of view become more important comparing to decisions about their inclusion to or exclusion from a certain course or program. Tomlinson (2012) considered evaluation as an activity done to make a judgment on the possible impacts of certain materials on learners. Evaluation has various types such as formative, summative, long term and short term (Nation and Macalister, 2010).The stance on evaluation can also be aggregated from qualitative evaluators, quantitative evaluators and mixed methods evaluators. As a result, there should be concord over systematized evaluation (Roberts, 1996). This systematic approach could be achieved by setting formal criteria for more consistent judgment.

Put differently, we need a tool to fetter the personal and biased judgment of the evaluators by expanding the agreed sets of criteria. The most effective

instrument for achieving systematic data has been the checklist (McGrath, 2002). Checklists can be developed in various models like close-ended or open-ended. This instrument has been mostly recommended by prominent scholars of curriculum design and development (Cunningsworth, 1995; Nation and Macalister, 2010), but the versatile nature of education possibly hinders the development of a checklist to satisfy all the possible needs. Hutchinson and Torres (1994) considered textbooks as agents of change; therefore, checklists ought to be dynamic and subject to change. In this respect, Anjaneyulu (2014) remarked that "ready-made checklists would fail to be used in various contexts and it is necessary to be modified and adapted for the purposes of their use" (p.185).

By scrutinizing a number of checklists available in the literature, it is recognizable that external and internal factors have mainly been under investigation (Cunningsworth, 1995; Daoud & Celce-Murica, 1979; Garinger, 2002; Nation and Macalister, 2010). External factors usually investigate the visuals and tangible parts of the textbooks and internal investigations mainly put the magnifier on the content of the textbooks. These are substantial areas, but some other jurisdictions are equally worthy of attention. The impacts of textbooks might go beyond the academic purposes toward ideological intentions.

In textbook evaluation, setting principle is of paramount importance. Tomlinson (2013) identified different theories with regard to teaching, learning and second language acquisition in which the mentioned lists of the principles were categorized into affective factors, wants, needs, experiential learning, etc. A dim expression of cultural awareness and personal voice has also been recognized, yet Tomlinson (2013) did not delve into the ideological aspects of textbook evaluation. Kumaravadivelu (2006) bolded that "No text is innocent" (p. 13); therefore, this statement is applicable to the discourses of textbooks. Developed skills and components of a language along with the activities and tasks in the textbooks may be ideologically laden. Therefore, oversimplified considerations of content, presentation, sequencing, goals, and formats do not suffice. There is a need for a tool to fortify against possible hidden ideologies, stereotyping and cross-cultural issues included in textbooks. Through reviewing different standardized

checklists available in the literature, it is recognizable that very limited attempts have been made to design a checklist from the lens of critical language studies.

Different approaches of critical language studies might have various macro intentions, but they share common goals namely emancipation and empowerment of learners. According to van Dijk (2001), routine discourse such as textbooks might shape our understanding of the world, our attitudes and values. By neglecting the ideological and cultural aspects of the textbooks, hidden thoughts and intentions of the authors and publishers might manipulate learners in their educational lifespan. Accordingly, Koupaee Dar, Rahimi and Shams (2010) claimed that by the absence of the critical stance on education "dominant group can win the thoughts and the hearts of the educated group of society who have vital roles in their own country" (p. 458).

The main goal of this study was to introduce a model to tackle the angles of English textbooks from the critical lens. It aimed at developing a checklist to investigate hidden thoughts and ideologies like racism, sexism, hegemony, marginalization, etc. The developed model is a dynamic checklist which can be adapted or modified for different educational intentions in various situations. It is worthwhile to mention that the nature of this checklist is supplementary which could be utilized besides the conventional checklists available or as a separate instrument for evaluating textbooks critically.

Conceptualizing criticality in critical language studies

Language can be viewed from different conceptual and perceptual perspectives. Kumaravadivelu (2006) theoretically conceptualized language as system, discourse and ideology. The third classification, language as ideology, has considerable features in common with the elaboration of Fairclough (1989) on the concept of critical language study. Accordingly, the analysis of the nexus between language, power and ideology is central in this approach. In other words, the political aspects of language which may naïvely manipulate the educator are an upfront action of the critique. In evaluation, the relationships between language and power can be investigated in the discourse and the practice of textbooks. Power can form and reform different types of discourse in professional and daily lives of

people (Fairclough, 1995). In this respect, identification of the sources of power and the possible manipulative intentions are worthy of attention for the future of education. This manipulation can have various sources for which ideology plays a significant role. Ideology could be located in forms, meanings and events in the texts (Fairclough, 1995), and all three factors ought to be included in the process of unveiling the ideological assumptions, hegemony and power relations in the texts and textbooks. To this end, the evaluators could probe the support of power and the dominant ideology in the process of developing texts and textbooks.

On the other hand, language and power can tackle the practice or the actual classroom teaching. Prettyman (2006) believed that the hierarchy of power in the classroom in which teacher was the source of knowledge and learners listening object could fetter learning. This approach considers learners as "resisters, receptacles, raw materials" (Williams and Burden, 1997, p.57). These considerations produce the situation for oppressing learners through force instead of creating possibility to learn languages.

On the contrary, transferring the power to the learners by providing circumstances for decision making could be helpful. In this case, learners and teachers become partners and equal for the purpose of learning. Kumaravadivelu (2006) believed that the power could be equally distributed to the learners, and they could get involved in decision making procedure through practicing "learn to learn" and "learn to liberate" (p. 176). In this regard, two intimidating political forces in applied linguistics which may hinder learning to liberate are industrialism and colonialism, which lead to pluralistic standards and standardization of language teaching and leaning (Pennycook, 1990). The process of standardization is interwoven with positivism and the concept of true knowledge or legitimate knowledge. The contrast becomes apparent when different schools of thought perceive knowledge differently. Knowledge from the positivistic perspective is in contrast to characteristics of Freire's (1970) problem-posing model; in which, knowledge is dynamic and dialogic.

In general, developing a textbook relies frequently on authors and institutions' beliefs, attitudes and their definition of knowledge and education. Dewey (1916) put true knowledge against practice and this

conflict may manipulate textbook development. In this respect, Freire (1970) believes that the banking system of education "achieves neither true knowledge nor true culture" (p.78). In addition, Kincheloe (2008) made a note of caution for instructors and learners by stating that

what we label as knowledge, the way it is arranged and presented, the ways it is taught and learned, and what is considered an appropriate display of having learned it is inseparable from the way we view the world, the purposes of education, the nature of good society and the working of the human minds (p.13).

Accordingly, critical pedagogy surpasses the reductionist, decontextualized, fragmented belief that knowledge could be solely measured. Conscious attention toward language teaching/learning is supported in critical pedagogy, in which not only form and meaning but also other social, cultural, economy and political aspects of language education can be questioned.

It is worthwhile to mention that the intuitive and anti-mentalist approaches to teaching/learning languages may pave the road for intended ideologies of the sources of power. Teachers and learners become servants of the monitored system rather than democratic educators. In this case, memorized chunks of knowledge (standardized) with no place for reflection are considered equal to learning true knowledge and culture. To illustrate, the term 'Policeman' (instead of police officer) may exist in the textbooks or may be used by teachers and learners without awareness of its sexist background. In the absence of critical investigations, the usage of this terminology may become legitimate as long as it is used phonologically, syntactically and semantically correct.

In applied linguistics, critical approaches connect language related issues "to questions of gender, class, sexuality, race, ethnicity, culture, identity, politics, ideology and discourse" (Pennycook, 2010, p. 4). In general, education cannot be viewed as an apolitical or a neutral entity, yet it seems that the political aspects of education are denied (Benesch, 1993). Hence, the denied aspects in language related studies are worthy of attention for textbook evaluation and classroom procedures.

Method

This enquiry was carried out to develop a critical checklist to fortify against possible hegemonic, ideological and manipulative intentions in global/commercial textbooks. However, providing a model in a labyrinth of definitions and intentions of critical studies besides the dynamic nature of EFL and ESL classrooms needs a framework to merge these two notions. To this end, classified areas of critical language studies by Pennycook (2010) were used in relation to the framework proposed by McDonough, Shaw and Masuhara (2013) to systematically consider these areas.

Participants

There are different views and debate about the sample size in factor analysis. According to MacCallum, Widaman, Zhang, and Hong (1999) communalities play more important role comparing to the sample size; on the contrary, others emphasized the large sample size to avoid the miss interpretation of correlation coefficient among samples (Field, 2009). To stay on the safe side, we decided to follow Field's (2009) suggestion with regard to sample size. The participants of this study included 230 MA students, 17 MA holders, 63 PhD candidates and 16 faculty members in Ferdowsi University of Mashhad, University of Mazandaran, Islamic Azad University of Qaemshahr, Islamic Azad University of Mashhad, and Tabaran Institute. Their major was applied linguistics, and they were aware of critical language studies. Moreover, their pedagogic experience was from 3 to 20 years (*M=9)*.There were five experts in the domain of critical language studies among professors, and the rest of the participants were completely familiar with this notion and its theoretical background. It is of value to mention that the targeted population (*t= 600)* was much more than the participants and some of the exclusions were due to unwillingness, unfamiliarity with the critical domain, careless completion, and uncompleted questionnaires. Therefore, the total number was truncated into 326 (*208 female and 118 male)* based on convenience sampling. The participants' age ranged from 25 to 54 with Persian as their first language.

Procedure for Exploring the Components of the Critical Checklist

The procedure of this study was constructed on several steps and stages for amassing, organizing and examining the items. According to Dornyei (2007), for developing a questionnaire, initial item pool is highly important. The first step included comprehensive review of the related literature which resulted in the initial model with 270 items. After a comprehensive investigation by colleagues for overlaps and repeated statements, 30 items were excluded. In the second phase of this step, the items were categorized based on their relevance to critical domains and areas of English programs (see Table 1).   Since the checklist is developed for textbook evaluation purposes, the content is of paramount importance. The category of content includes critical investigation in listening, speaking, writing and reading besides components like grammar, pronunciation and visuals.

Table 1.
The categories of the developed checklist

Categories	Critical domains	Targeted areas of English program
Curriculum/syllabus Attributions	CALP	Curriculum/syllabus
Methodological barriers	CP	Methodology
Learners and teachers' constraints	CLA-CCDA	Learners and teachers
Content (Artwork, skills and components)	CDA	Materials

The second stage was semi-structured interview of experts (16 faculty members of the mentioned universities) to accumulate extra information and to check the clarity and appropriateness of the items. The expert review could be considered as a testing process for the tentative draft (Giesen, Vis-Visschers and Beukenhorst, 2012). Accordingly, the experts' views were sought after their review of the initial draft of the checklist via a thirty-minute interview. This interview was an endeavor to obtain confirmation over the first draft of the checklist and to explore further areas for inclusion. Most of the interviewees were uninterested in tape recording; therefore, we decided to include two fast writers to participate in the interview sessions to

duplicate the utterances. Extremely valuable hints and comments were received from this stage. This emphasis adjoined one of the crucial elements of critical language studies to the items. With regard to the opinion about the checklist one of the repeatedly mentioned issues by the faculty members was the length. In this respect, the checklist reduced into 70 items for the validation purposes based on the expert judgment.

Following the standard procedure for developing a checklist (Dornyei, 2007), a 5-point Likert scale ranging from "unimportant" to "very important" was designed for the piloting stage. Then, the tentative checklist was administered to 30 PhD students and 5 professors of Islamic Azad University of Qaemshahr. The estimate of reliability for 70 items and 35 participants showed .874. According to Barker, Pistrang, and Elliott (1994), the desired threshold of reliability is above .70, which indicates acceptable estimation for this instrument. Then, the face to face method was used for the administration of the checklist to the targeted participants.

Design and Procedure

Exploratory factor analysis (EFA) and partial confirmatory factor analysis (PCFA) were used in this study via SPSS (version 17). In the initial stage of the exploratory factor analysis, descriptive statistic, the normality of the distribution was sought via skewness and kurtosis values. Next, principal components analysis was used as the method of component extraction with orthogonal (varimax) rotation. Then, the estimates of Cronbach's alpha for the components were sought.

PCFA comprises the second stage of the data analysis after component extraction. In an educative article about PCFA, Gignac (2009) elaborated the usefulness of this technique before confirmatory factor analysis. Accordingly, this study utilized this technique to check the possibility of close-fit of the components for the forthcoming studies. In this respect, SPSS (version 17) with maximum likelihood method was used to identify implied and null chi-square values and the degrees of freedom. These values were crucial in calculating incremental and absolute close-fit indexes namely normed fit index (NFI), Tucker-Lewis index (TLI), comparative fit index (CFI), root mean square error of approximation (RMSEA) and

standardized root mean residual (SRMR) (Gignac, 2009). All the calculations were manually done in this stage (for the formulas see Gignac, 2009).

Results

In the first stage of the analysis, statistics such as descriptive analysis, EFA along with internal consistency of the hypothetical model were brought into consideration. Furthermore, the appropriateness of the model for the subsequent confirmatory analysis was underpinned by PCFA in the second stage. Before the initial principle component analysis (henceforth PCA), a correlation analysis had to be conducted to indicate the possible problematic components (Field, 2009). Accordingly, 26 items with extremely high and low correlations were removed for the main study. Thus, 44 items were left for the main analysis (see Table 2).

Table 2.
Remained items of the critical checklist

N	Items
1	Reflective reading is encouraged in the textbook.
2	There are activities which promote reading between the lines.
3	The reading texts practice cultural marginalization.
4	The textbook encourages critical responses in writing skill.
5	The writing exercises allow critical discussion of the topic/idea.
6	The writing exercises allow the learners to express their own opinion with certainty.
7	The textbook relies on a universally dominated method.
8	The textbook restricts the teacher's adaptability for applying locally suitable methodology.
9	The textbook relies on a universally dominated method.
10	The quotes are ideologically laden.
11	The textbook includes racist language.
12	The textbook includes sexist language.
13	The textbook encourages critical thinking.
14	Controversial issues such as Aids, abortion, euthanasia, and as such are included in the content.
15	The content of the textbook includes famous people from all over the world.
16	The recommended syllabus exercises norms and values of English speaking societies.
17	The recommended syllabus (or advocated lesson plan) bans the first language/culture of the learners in English classrooms.
18	The recommended syllabus is developed based on ESL/EFL teachers' and learners political, economic and cultural constraints.
19	The recommended syllabus is developed according to limitations of English speaking

	societies.
20	The syllabus is directly influenced by publisher's intentions (political, economical, and cultural).
21	The textbook is designed to give identity to learners as speakers of the second language.
22	The textbook develops assertive behavior.
23	The textbook develops submissive behavior.
24	The textbook is designed to raise learners' awareness on social phenomena.
25	The textbook develops sexist vocabularies.(For example: policeman instead of police officer, businessman instead of business manager)
26	The textbook encourages different grammatical complexity for politeness in different social relations.
27	The textbook encourages equal identity for native and non-native teachers.
28	The teacher's guide encourages dialogic interaction between the teacher and learner.
29	The textbook makes teachers servant of the globally identified system.
30	The teacher's guide restricts teacher's autonomy.
31	The visuals include different races.
32	The textbook includes equal number of male and female in the pictures.
33	The sexist ideology is included in the pictures.
34	The visuals represent a utopian world.
35	Different aspects of reality (hunger, poverty, etc.) are depicted.
36	The visuals are appropriate for the exported country.
37	The listening sections familiarize learners with the distinctive accents of World Englishes.
38	The textbook promotes American or British accents only.
39	Different dialects are used in listening sections.
40	Particular news (with hidden intentions) is used in listening section.
41	The speaking tasks prioritize teachers' opinion as a source of power.
42	The speaking tasks encourage cooperation between silent and talkative learners.
43	The speakers are mostly men.
44	The activities in the textbook encourage each learner to freely express her or his own opinion in the classroom.

Descriptive Data Analysis

In the initial step of the analysis after checking for missing data, the normality of distribution was pursued. The values of skewness and kurtosis along with their standard errors for the variables were checked. All the values were within the range of±2; therefore, the findings revealed symmetric and normally distributed results. In addition, the visual histogram check revealed an acceptable bell-shaped normality among the components.

Exploratory Factor Analysis (EFA)

In the main phase of the data analysis, in order to determine statistical support for the components, PCA method was used for 44 variables with

orthogonal rotation (varimax). With regard to appropriateness of the sample size for the validity purposes in this checklist, the measure of Kaiser-Meyer-Olkin (KMO) was selected. The overall value of KMO turned out to be .876 (KMO= .876) which could be considered reasonable for sampling adequacy and appropriateness of factor analysis (Field, 2009). In addition, the diagonals ranged between .706 and .926 (*M*= .868*)* in anti-image matrices which are considerably higher than the cut-off value of 0.5 (Field, 2009).Therefore, KMO for individual items could also be considered reasonable. On the other hand, Bartlett's test of sphericity $\chi^2(946) =$ 10942.702, $P \leq 001$ indicated statistically significant result for the patterned correlations between the items (see Table 3).

Table 3.
Kaiser-Meyer-Olkin and Bartlett's Test

Tests		Results
KMO Measure of Sampling Adequacy.		.876
Bartlett's test of Sphericity	Approx. Chi-Square	10942.702
	Df	946
	Sig.	.000

   Preceding the main analysis, eigenvalue of 1 as Kaiser's criterion was selected for the extraction of the components. The result indicated that 10 components were above the point of inflexion with eigenvalue more than 1 (see Figure 1).
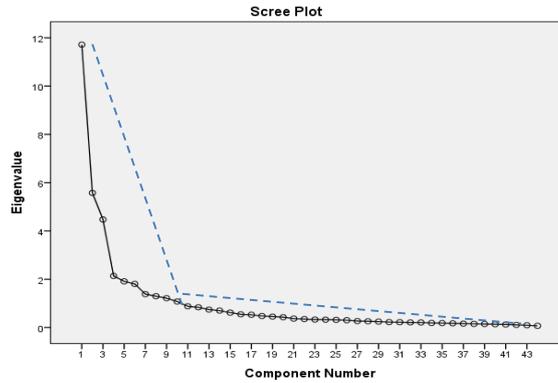
Figure 1. Scree plot indicating that the data have 10 components

The cross section of the horizontal and vertical dotted lines depicts the inflexion point. These statistically identified components were the pre-classified items which were developed based on the critical domains discussed above. The minimum factor loading value of this analysis was selected to be 0.4 based on Field's (2009) suggestion. The result of this analysis revealed eigenvalues of 4.16, 4.16, 4.10, 3.65, 3.29, 3.08, 2.76, 2.67, 2.50, and 2.18 respectively for the component 1 to 10 after rotation. Table 4 shows initial and rotated eigenvalues, variance and cumulative percentage before and after rotation.

Table 4.
Eigenvalues before and after rotation

	Initial Eigenvalues			Rotation Sums of Squared Loadings		
Component	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	11.71	26.63	26.63	4.16	9.47	9.47
2	5.57	12.67	39.30	4.16	9.46	18.93
3	4.47	10.17	49.47	4.10	9.31	28.25
4	2.13	4.85	54.32	3.65	8.31	36.56
5	1.91	4.34	58.67	3.29	7.48	44.05
6	1.80	4.10	62.77	3.08	7.01	51.06
7	1.38	3.14	65.92	2.76	6.28	57.34
8	1.29	2.95	68.87	2.67	6.08	63.42
9	1.21	2.76	71.63	2.50	5.69	69.12
10	1.07	2.45	74.09	2.18	4.97	74.09

These findings along with the scree test indicate that all 10 components could be extracted. This number was retained for factor loadings after rotation. Factor loadings constituted the final step in PCA which according to Field (2009), "were a gauge of substantive importance of a given variable to a given factor" (p.644).   Table 5 shows the loaded components after rotation. The components were labeled based on the experts' judgments and the criteria which could be investigated in English language programs. The items that clustered on the same components suggested syllabus attributions (S), methodological barriers (M), content (C), listening and speaking (L/S), reading and writing (R/W), artwork and utilitarian attributes (A), vocabulary and grammar (V/G), pronunciation (P), learners' considerations (L), and teachers' considerations (T).

Table 5.
The results of EFA (N=326)

Exploratory Factor Analysis

	Components	C	L/S	R/W	A	M	L	T	S	P	V/G
Items											
1	R/W			.861							
2	R/W			.868							
3	R/W			.816							
4	R/W			.887							
5	R/W			.846							
6	R/W			.500		.630					
7	M					.840					
8	M					.894					
9	M					.852					
10	C	.411				.654					
11	C	.824									
12	C	.916									
13	C	.911									
14	C	.883									
15	C	.650									
16	S	.422							.515		
17	S								.805		
18	S								.823		
19	S								.749		
20	S								.402		
21	L						.647				

Item						
22	L			.790			
23	L			.719			
24	L			.613			.422
25	V/G						.786
26	V/G						.859
27	T				.635		.453
28	T				.763		
29	T				.782		
30	T				.567		
31	A		.559				
32	A		.781				
33	A		.746				
34	A		.791				
35	A		.762				
36	A		.416			.715	
37	P					.764	
38	P					.752	
39	L/S	.752					
40	L/S	.754					
41	L/S	.766					
42	L/S	.772					
43	L/S	.588					
44	L/S	.512					

Extraction Method: Principal Component Analysis.
Rotation Method: Varimax with Kaiser Normalization.

There are 6 items with simultaneous loading on different factors in the Table 5. Accordingly, item 6, '*The writing exercises allow the learners to express their own opinion with certainty.*' had simultaneous loadings (.50-.63) on the R/W and M. Item 10, '*The quotes are ideologically laden*' had factor loadings (.41-.65) on C and M. Also, item 16, '*The recommended syllabus exercises norms and values of English speaking societies*' had factor loadings (.42-.51) on both C and S. Item 24, '*The textbook is designed to raise learners' awareness on social phenomena*' had factor loadings (.61-.42) on L and V/G. Item 27, '*The textbook merely provides non-native teachers the identity (art of teaching rather than filling students with materials) equal to native teachers*' revealed simultaneous loading (.63-.45) on T and V/G. Finally, item 36, '*Visuals are appropriate for the exported*

country' had component loadings (.41-.71) on A and P. These items were excluded from the rest of the study and the list of the tentative items truncated into 38.

Internal Consistency

To indicate that items consistently reflect their components or their constructs, Cronbach's alpha was utilized for each component.

Table 6.
Reliability for each component

	C	L/S	R/W	A	M	L	T	S	P	V/G
N	5	6	5	5	3	3	3	4	2	2
Alpha	.910	.883	.920	.888	.898	.848	.853	.812	.834	.874

The result of Cronbach's alpha in Table 6 indicated that reliability estimate of all the 10 components (38 items) were above .8, ranging between .81 and .92. Based on Barker et al. (1994) threshold, the result of alpha indicated good internal consistency for the items and the corresponding components.

Partial Confirmatory Factor Analysis

In the second stage of this study, the maximum likelihood method with orthogonal (varimax) rotation was used to identify crucial elements for close-fit index. The maximum likelihood chi-square value was equal to 960.31 with 368 degrees of freedom ($P \leq .001$). This value was significantly smaller than the null value 9030.84 with 703 degrees of freedom ($P \leq .001$). These values were crucial for calculating the close-fit indexes (Gignac, 2009).

Table 7.
The results of NFI, TLI, CFI, RMSEA, SRMR

NFI	TLI	CFI	RMSEA	SRMR	$\chi^2 n$	$df\,n$	$\chi^2 i$	$df\,i$
0.89	**0.86**	**0.96**	**0.05**	**0.04**	9030.84	703	960.31	368

According to Hu and Bentler (1999) the obtained values of NFI, TLI, and CFI should be close to cutoff value of .95 to represent a good fit between tentative model and the data. The result for incremental indexes in the Table 7 revealed lower than acceptable values (NFI, TLI, $\leq$ .95). Among these

indexes, CFI passed the cutoff value of .95 (see Table 7).The absolute close-fit index of RMSEA should be close to cutoff value of .06 (Hu and Bentler, 1999). In accordance, the result indicated that the value of RMSEA was lower than the acceptable value (RMSEA≤ .06). Moreover, the value of SRMR also could not reach the cutoff score of .08(SRMR ≤.08). Therefore, the hypothetical model could not be considered acceptable for confirmatory factor analysis, as the overall results (except for CFI) of incremental and absolute close-fit indexes were not standard.

Discussion

The present study explored 10 critical components that are believed by the authors to have the potentials to fortify against hegemonic, unequal, unfair and unjust characteristics which could be transmitted by the imported commercial textbooks. These 10 components include 38 items developed for different areas of commercially designed textbooks. The component of methodological barriers is inline with Kumaravadivelu's (2006) postmethod pedagogy and Friere's (1970) problem-posing model of education. Put differently, this component makes textbook evaluators aware of pedagogical and ideological forces from global to local settings. This component tries to tackle power and domination of western theorists for the rest of the world with regard to teaching methodology advocated in commercial textbooks. The component of syllabus attributions investigates the hidden curriculum and hidden policy for spreading the ideology of monolingualism. In other words, the identity of English language learners is directly related to their linguistic assimilation. This component is inline with the critical approach to language policy (Phillipson, 1992), and the concept of hidden curriculum in which the sources of power control minorities' thoughts (Alimorad, 2016). This component is so vast that it cannot be evaluated in small scale, yet it can awaken evaluators about different political, economic and cultural outcomes which a textbook might have. In other words, the recommended syllabus or the progress chart may draw the attention of teachers to the surface rather than deep constructions of materials, but this component has a critically enlightening nature. Roohani and Heidari (2012) found three main texts about genius male and one text (in the workbook) about genius woman in a commercially developed textbook. In the case of teaching such books, teachers or textbook evaluators may skip the forms of inequality since their attention is toward depositing knowledge.

Learner considerations, in line with concepts of empowerment and emancipation, are rooted in critical language awareness (see Clark, 1992). Accordingly, the items check the voices of learners in relation to the learning of English language. This component investigates learners' responses to tasks and lessons. The negotiation of identities (suppressed or critical) which commercially designed textbooks encourage among learners is under the magnifier (Taylor, Despagne and Faez (2018).The other component, teacher considerations, tries to evaluate teachers' maneuverability in the classroom. This component consists of three critical items in line with Kumaravadivelu's (2006) postmethod teacher and Friere's (1970) concept of dialogical teaching. In fact, this component not only evaluates teacher's autonomy for practicing local theories but also considers the power relation between teachers and learners.

In addition, content is one of the core components of this checklist. Nation and Macalister (2010) stated that the focus of language lessons could be on "Language, Ideas, Skills or Text (Discourse)" (P.71). Accordingly, the component of content in this checklist focuses on discourses of commercially developed textbooks. This component is derived from Fairclough's (1989) CDA model to highlight the possibility of inequality, injustice, and political forces in the discourse of such textbooks. It is worth mentioning that the components of artwork listening and speaking, reading and writing, artwork and utilitarian attributes, vocabulary and grammar, and pronunciation could be considered as the sub-components of the content. To elaborate, the component of V/G remains the shared element between content of the textbooks (see Nation and Macalister, 2010) and Fairclough's (1989) three dimensional model of critical discourse analysis. According to Widdoson (2004) discourse was unrestricted to written or spoken forms, and it included other elements such as visuals. The component of artwork critically investigates the pictures of commercially designed textbooks, which goes beyond the ordinary pictorial investigation like color drawings or color photos (see Hill, 2013). In essence, this component makes the evaluators aware of sexist, racist and dehumanized presentation of the world in the pictures. For instance, the frequent use of men in firefighting and women in nursing can be considered as an illustration for a sexist ideology in visuals.

These critical components were explored in EFA stage of this study. To investigate the possibility for further study with regard to validation process, PCFA was conducted. However, the result of PCFA was unsatisfactory. It is of value to mention that the present study followed a rigid factor analysis. In the literature some studies subjectively did not eliminate the items with

simultaneous loadings, but the cross-loadings were excluded in this study for PCFA.

Conclusion

This study explored 10 critical components including 38 items for evaluation of textbooks. The checklist is derived from the critical domains categorized by Pennycook (2010) namely CALP, CP, CLA-CCDA, and CDA. These areas are in line with practical aspects of English language programs with focal attention toward actual classroom procedures. In other words, the explored model tries to fortify against ideological manipulations for learners and teachers in English language classrooms using commercially developed textbook. This critical checklist could be used as a supplement besides other checklists (attentive to language as system and language as discourse). Textbook evaluators, curriculum designers, and teachers can use this critical checklist to identify some of the obstacles in the path of democratic education.

This study was carried out in Iran (Mazandaran and Razavi Khorasan provinces) and the participants academic levels were MAs (MA students and MA holders) and PhDs (PhD candidates and PhD holders). It is recommended that this study be replicated in other contexts. In addition, the sample size of this enquiry was 326; thus, broader samples are also recommended for further studies on this checklist. Other critical domains could also be utilized to cover wider areas in the realm of TEFL. It is highly recommended to develop items for different problematic areas which evaluators might find productive in their settings.

References

Alimorad, Z. (2016). Cultural values represented in first certificate masterclass taught in Iran: Uncovering the hidden curriculum. *Sage, 6*(1), 1-8.

Allwright, R. (1981). What do we want teaching materials for? ELT Journal, 36, 5-18.

Anjaneyulu, T. (2014). A critical analysis of the English language textbooks in Andhra Pradesh, India. *ELT Research Journal, 3*(4), 181-200.

Barker, C., Pistrang, N., & Elliot, R. . (1994). *Research methods in clinical and counselling psychology.* Oxford: John Wiley & Sons.

Benesch, S. (1993). ESL, ideology, and politics of pragmatism . *TESOL Quarterly* , 705-716.

Clark, R. (1992). Principles and practice of CLA in the classroom. In N. Fairclough (Ed.), *Critical language awareness* (pp. 117-140). London: Pearson Education.

Cunningsworth, A. (1995). *Choosing your coursebook.* Oxford: Heinemann.

Daoud, A., & Clce-Murcia, M. (1979). Selecting and evaluating a textbook. In M. &. Clce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 302-207). Cambridge : MA: New bury House Publishers.

Davies, A. (2006). What do learners really want from their EFL courses? *ELT Journal, 60*(1), 1-12.

Dewey, J. (1916). *Democracy and education: An introduction to the philosophy of education.* New York: Macmillan.

Dornyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative and mixed methodologies.* Oxford : Oxford University Press.

Edge, J. & Wharton, S. (1998). Autonomy and development: Living in the material world. In B. Tomlinson (Ed.), *Materials development for language teaching* (pp. 295-310). Cambridge: Cambridge University Press.

Fairclough, N. (1989). *Language and power.* New York: Longman.

Fairclough, N. (1995). *Critical discourse analysis: The critical study of language* (2nd ed.).
London: Longman.

Field, A. (2009). *Discovering statistics using SPSS.* London : Sage.

Freire, P. (1970). *Pedagogy of the oppressed.* New York: The Continuum International Publishing.

Garinger, D. (2002). *textbook selection for the ESL classroom.* Eric Digest. Retrieved November   24, 2016 from http://www.cal.org.

Giesen, D., Vis-Visschers, R., & Beukenhorst, D. (2012). *Questionnaire development.* Hague: Statistics Netherlands.

Gignac, G. E. (2009). Partial confirmatory factor analysis: Describe and illustrated on the NEO-PI-R. *Journal of Personality Assessment, 91*(1), 40-47.

Hadley, G. (2014). Global textbooks in local context: An empirical investigation of effectiveness. In N. Harwood (Ed.), *English language teaching textbooks: Content, consumption, production* (pp. 205-238). London: Routledge.

Hill, D. A. (2013). The visual elements in EFL coursebooks. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 157-166). London: Bloomsbury.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexesin covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1-55.

Hutchinson, T. & Torres, E. (1994). The textbook as agent of change. *ELT Journal , 48* (4), 315-328.

Koupaee Dar, Z., Rahimi, A., & Shams, M. R. (2010). (2010). Teaching reading with a critical attitude:Using critical discourse analysis (CDA) to raise EFL university students' critical language awareness (CLA). International Journal of Criminology and Sociological Theory, 3(2), 457-467.Kincheloe, L. J. (2008). Critical pedagogy and the knowledge wars of the twenty- first century. *International Journal of Critical Pedagogy, 1*(1), 1-22.

Kincheloe, L. J. (2008). Critical pedagogy and the knowledge wars of the twenty- first century. *International Journal of Critical Pedagogy, 1*(1), 1-22.

Kumaravadivelu, B. (2006). *Understanding language teaching:From method to postmethod.* New Jersey: Lawrence Erlbaum Associates. MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99.

MacCallum, R. C., Widaman, K. F., Zhang, S., & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods, 4*(1), 84-99.

McDonough, J., Shaw, C., & Masuhara,. (2013). *Materials and Methods in ELT: A teacher's guide.* Oxford: Wiley-Blackwell.

McGrath, I. (2002). *Materials evaluation and design for language teaching.* Edinburgh: Edinburgh University Press.

Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design .* New York : Routledge.

O'Neil, R. (1982). Why use textbooks? *ELT Journal , 36* (2), 104-111.

Pennycook, A. (1990). Towards a critical applied linguistics for the 1990s. *Issues in Applied Linguistics, 1*(1), 8-28.

Pennycook, A. (2010). Critical and alternative directions in applied linguistics . *Australian Review of Applied Linguitsics*, 1-16.

Phillipson, R. (1992). *Linguistic imperialism.* Oxford: Oxford University Press.

Prettyman, K. (2006). Rethinking power: Learners as teachers and teachers as learners. In C. Dozier, P. Johnston, & R. Rogers, *Critical literacy/critical teaching: Tools for preparing responsive teachers* (pp. 46-53). New York: Teachers College Press.

Roberts, J. T. (1996). Demystifying materials evaluation. *Pergamon , 24* (3), 375-389.

Roohani, A., & Heidari, N. (2012). Evaluating an instructional textbook: A critical discourse perspective. *Issues in Langauge Teaching, 1*(1), 123-158.

Taylor, S. K., Despagne, C., & Faez, F. (2018). Critical language awarness. In J. I. Liontas (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1-16). Hoboken: NJ: Wiley.

Tomlinson, B. (2012). Materials development for language learning and teaching. *Language Teaching, 45*(2), 143-179.

Tomlinson, B. (2013). Materials evaluation. In B. Tomlinson (Ed.), *Developing materials for language teaching* (pp. 21-43). London: Bloomsbury.

Ur, P. (1996). *Acourse in language teaching: Practice and theory* . UK: Cambridge University Press.

van Dijk, T. A. (2001). Critical discourse analysis. In D. Tannen, D. Schiffrin, & H. & Hamilton (Eds.), *The handbook of discourse analysis* (pp. 352-371). Oxford: Blackwell.

Widdowson, H. G. (2004). *Text, Context, Pretext: Critical issues in discourse analysis.* Oxford: Blackwell Publishing.

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

Biodata

Siavash Zokaeieh is a Ph.D. candidate at Islamic Azad University, Qaemshahr Branch, Iran. His main reseach area is critical language studies.

Shaban Najafi Karimi is an assistant professor at Islamic Azad University, Qaemshahr Branch.His main research interests include teaching langauge skills and materials development.

Farzaneh Nouri is currently a Ph.D. candidate in TEFL at Islamic Azad University, Qaemshahr Branch, Iran. Her research interest is critical language testing.

Afsaneh Hakimzadeh is a Ph.D. candidate in TEFL at Islamic Azad University of Qaemshahr, Iran. Her research interests is checklist development in postmethod pedagogy.