

Criteria in English language textbook evaluation checklists: A systematic review

Mohamad Syafiq Ya Shak^{1,2}

Intan Safinas Mohd Ariff Albakri²

Siti Shuhaida Shukor²

Mohd Haniff Mohd Tahir²

¹Academy of Language Studies, Universiti Teknologi MARA Perak Branch, Perak, Malaysia

²Faculty of Languages and Communication, Universiti Pendidikan Sultan Idris, Perak, Malaysia

Abstract

A systematic review was carried out to examine the evaluation criteria in textbook evaluation checklists for evaluating ESL/EFL textbooks worldwide and to identify gaps and additional criteria that could be included in future textbook evaluation checklists in order to meet the demands of current teaching and learning situations. Two databases – Scopus and the Web of Science – were explored to collect data. Primary searches between 2011 and 2021 revealed 92 studies on the topic under investigation. After scrutinising abstracts and removing duplicates, 36 studies were retained for further analysis. A thematic analysis was conducted to derive themes for the criteria enlisted in these studies. The themes of criteria that emerged were: (1) practical considerations; (2) layout and design; (3) language skills; (4) language activities and tasks; (5) topic/subject of the content; (6) appropriateness for students; (7) cultural considerations; (8) supplementary materials; and (9) alignment with the language programme's aims and objectives. It is recommended that future textbook evaluation checklists focus on criteria that relate closely to ESL/EFL textbook users, especially in terms of cultural representation, promoting self-study, and technology integration, especially in online distance learning.

Keywords: checklist, EFL, ELT, ESL, textbook evaluation

Introduction

Textbook evaluation can be carried out through various approaches. A criterion-based checklist is one of the most popular (Abdel Wahab, 2013; McGrath, 2016; Richards, 2016). In the context of English language teaching (ELT), Mukundan et al. (2011) define a checklist as an instrument that helps practitioners to evaluate materials such as textbooks for English language learning. According to Brown (2001), textbook evaluation checklists consist of a comprehensive list of criteria that allows the evaluation process to be completed systematically. These criteria could include aspects such as a textbook's physical appearance, its tasks, exercises, and activities, coverage of language skills (reading, writing, listening, speaking, grammar, vocabulary), its relation to a syllabus and curriculum, and its compatibility with learners (see Appendix 1).

There are many reasons why ELT textbook evaluators worldwide use evaluation checklists. For instance, they enable detailed and in-depth evaluation, especially if qualitative measures are used (Cunningsworth, 1995; Skierso, 1991; Mukundan et al., 2011; Demir & Ertas, 2014). A checklist can also be easily replicated (Ellis, 1997) and customised to suit the needs of future users (Mukundan & Ahour, 2010). Lastly, it is also believed to be economical, enabling much information to be recorded in a relatively short time (McGrath, 2016).

Evaluating a textbook using a checklist or other suitable evaluation tool has become more significant due to the recent COVID-19 pandemic. With prolonged disruptions to usual teaching and learning, online distance learning activities have steadily gained prominence as schools closed. There were concerns about digital access among students, especially in developing countries like Malaysia and other Southeast Asian countries (Kapasia et al., 2020; Lau, 2020; Li & Lalani, 2020; Upadhyaya et al., 2020; Shak et al., 2021b). Some students had inadequate access to online learning sessions (Di Pietro et al., 2020; United Nations, 2020; Wan, 2020). There were also disparities in terms of internet connectivity, especially between students in urban and rural areas (Arumugam, 2020; Kapasia et al., 2020; Lim, 2020; Yee, 2020) and between those with different socioeconomic backgrounds (Andrew et al., 2020; Kapasia et al., 2020; Lau, 2020; Li & Lalani, 2020; Upadhyaya et al., 2020).

Chabbott and Sinclair (2020) argued that printed textbooks could help address some of the issues related to online distance learning, especially for students from lower- to middle-income families. Numerous researchers highlighted that limited digital devices and insufficient access to the internet due to the geographical location are some of the limitations students face (Chabbott & Sinclair, 2020; Bell et al., 2020; Buchbinder, 2020). Textbooks are convenient for students (Millar & Schrier, 2015), they offer the cheapest way of providing learning materials (Ur, 1996), and they do not have to be accessed through an expensive device or software (Engbrecht, 2018). In addition, in online distance learning, textbooks can be used by teachers and caregivers through phone calls or messages to provide instruction and guidance to caregivers (Chabbott & Sinclair, 2020). Most importantly, a printed textbook supports self-directed learning (Cunningsworth, 1995; Rubdy, 2013; Ahmed, Yaqoob, & Yaqoob, 2015), in which learning can take place beyond the classroom and at home (Jusuf, 2018). Ur (1996) stated that students could use a textbook to study as well as review and monitor progress, and she claimed that learners without printed textbooks are more dependent on teachers. However, most textbooks were not produced or designed to be used during a pandemic, where the learning occurred remotely and online. They were mainly designed to support in-person learning at schools.

In the context of textbook evaluation checklists, Roberts (1996), Byrd (2001), and McDonough et al. (2013) argued that different instructional settings would require a different set of textbook evaluation criteria. Hence, in current teaching and learning situations, which rely heavily on technological advances and students' ability to learn independently, the criteria employed to evaluate textbooks may need to be revised, and new criteria may need to be added to make the checklist more comprehensive.

Therefore, this review seeks to answer the following questions: (i) What criteria in textbook evaluation checklists have been used to evaluate ESL/EFL textbooks around the world in the past decade? (ii) What gaps and additional criteria could be included in future textbook evaluation checklists to meet the demand of the current teaching and learning situations?

Methodology

This study employed a systematic review. This is a scientific method for gathering insights into a specific research domain while aiding future studies in identifying gaps and trends in previous and current studies (Majid & Salam, 2021). In completing this systematic review, the authors adopted a publication

standard suggested by Moher et al. (2015), called Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). Using the protocols proposed by PRISMA, the authors explained the inclusion and exclusion criteria in choosing the previous studies for this systematic review, the sources of these studies, and how the data for this systematic review was collected and synthesised.

Inclusion / exclusion criteria

The researchers used several attributes in selecting the published studies for this systematic review. The selected articles were:

- published from 2011 to 2021
- journal articles with empirical data and not review or conceptual papers
- full papers, not just abstracts
- written in English
- on the evaluation of physical / printed textbooks
- on English language textbooks, whether they are produced locally or imported from native-speaking countries

Development of a web search strategy

Five main keywords were identified, based on the formulated research questions: textbook, evaluation, checklist, ESL, and EFL. The researchers also sought synonyms, related terms, and variations to enrich these keywords. The combinations of these keywords and their variations were processed by using search functions such as field code functions, phrase searching, wildcards, truncation, and Boolean operators in two databases, Scopus and Web of Science (WOS). Samsuddin et al. (2021) suggested that these databases were chosen due to their leading citation indexing systems and controlled article quality. In addition, as suggested by Tamilchelvan & Rashid (2017), the main search strategy was to consider the title, abstract, and key terms in using these search functions. Table 1 shows the search strings used in these two databases:

Table 1. Search string used in the selected databases

Database	Search String
Scopus	TITLE-ABS-KEY (((textbook) AND (checklist) AND (evaluat* OR analys* OR analyz*) AND (e*1)))
Web of Science	TOPIC: (((textbook) AND (checklist) AND (evaluat* OR analys* OR analyz*) AND (E*L)))

Observation protocol

After all the publications within the stated search parameters had been gathered, the titles of these publications were scrutinised before the abstracts were vetted for adequacy and relevance. If the information in the abstract was deemed insufficient, the whole publication was screened to determine whether it fits the selection criteria of this study.

In the initial stage, 105 publications were gathered from Scopus and Web of Science databases, based on the search strings in Table 1. Then, 18 publications were removed based on the publication titles, leaving 87 for further screening. Next, 13 publications not published between 2011 and 2020 were excluded, leaving 74 publications for the next stage. Next, one more publication was excluded because only an abstract was published. Three more were rejected because they were review and conceptual papers and not full journal articles with empirical data, leaving 70 articles to be further scrutinised.

Three of these 70 articles were excluded as they were not written in English, and two more were rejected as they did not evaluate a physical printed textbook (they focused on eBooks), leaving 65 articles for the final screening. In this final stage, 29 articles not related to teaching English as a Second Language (ESL) or English as a Foreign Language (EFL) were excluded after the articles were screened in full, leaving a final 36 articles to be reviewed. Table 2 summarises the observation protocol for the articles chosen for this study:

Table 2. Exclusion criteria in selecting articles to be reviewed

Exclusion Criteria	Excluded articles	Remaining articles
*Duplication between WoS and Scopus databases	18	87 (from the initial 105)
Not published between 2011 and 2021	13	74
Not a full publication	1	73
Not a journal article paper	3	70
Not in English	3	67
Do not evaluate a physical textbook.	2	65
Not related to the teaching of English as a Second or Foreign Language (ESL/EFL)	29	36 (final number to be reviewed)

Data extraction and analyses

The selected articles were thematically analysed. This is a form of analysis where similar patterns, themes, and relationships within a particular set of data are identified, analysed, and reported (Braun & Clarke, 2019). It is also ‘an appropriate method of analysis for seeking to understand experiences, thoughts, or behaviors across a data set’ (Kiger & Varpio, 2020, p.846).

In completing the thematic analysis for this review, the researchers followed the steps suggested by Kiger and Varpio (2020). First, the researchers read the selected articles more than once to familiarise themselves with the content. Next, the researchers generated initial codes from the reading process and extracted any data related to the research questions. Then, the themes related to textbook evaluation criteria were generated inductively, where the researchers tried to observe any interests, similarities, and connections in the extracted data. The themes were then discussed with an intercoder, a senior researcher in the field, who agreed that textbook evaluation criteria had overlapping themes in the selected studies. This helped the researchers streamline the major themes of textbook evaluation criteria before data was added.

Findings

Context of the selected studies

The country where the study was conducted

Of the selected 36 studies, interestingly, about 60 percent (21) were completed in Iran. Apart from that, four studies were done in Spain, two in Malaysia, and one in Afghanistan, Algeria, Chile, Indonesia, Morocco, Pakistan, Saudi Arabia, South Korea, and Turkey. (See Appendix 2 for the complete list.)

The instrument used for data collection – adapted or self-developed checklist

Regarding the instrument used for data collection in the 36 selected studies, 7 percent (27 studies) employed textbook evaluation checklists adapted from other evaluators and previous studies, while 25 percent (9 studies) developed their own checklists. The most adapted textbook evaluation checklist was that of Litz (2005), used by six studies, followed by that of Byram (1993), used by four. The third most adapted textbook evaluation checklist was jointly that of Bloom (2001) and Miekley (2005), each adapted in two studies. Additionally, twelve studies adapted their checklists from twelve different sources. On the other hand, nine studies developed their own checklists to complete their textbook evaluation exercises. (See Appendix 3.)

Participants in the textbook evaluation exercises

In terms of the participants in the studies – those who conducted the textbook evaluation exercises – most of the evaluation exercises (48 percent or 17 studies) were completed by the researchers themselves. Moreover, seven of these studies employed teachers and students as evaluators; six used teachers as evaluators, while only one employed students using the evaluated textbook as evaluators. In addition, one study each utilised teachers, postgraduate students, and teacher trainees, respectively, as evaluators. In addition, English teaching experts were invited as evaluators in three of these studies. (See Appendix 4.)

The focus of the selected studies

Almost half of the studies (17) evaluated English language textbooks in general, including aspects such as physical appearance, content, language skills, language tasks, exercises, and activities, practicality, topics, and supplementary materials to determine the textbook's effectiveness and usability. (See Appendix 5.)

Apart from that, seven studies focused on various facets of cultural content and representation, such as a textbook's cultural exophoric references (Azadsarv et al., 2015), positive, negative, and neutral cultural representations (Abd Rashid & Engku Ibrahim, 2017), national, international, and target culture content (Abbasian & Biria, 2017), cultural aspects (Raigon-Rodriguez, 2018), culture types (Larrea-Espinosa & Raigon-Rodriguez, 2019), intercultural perspectives (Amerian & Tajabadi, 2020), and cultural dimensions (Ariawan, 2020).

In addition, English language skills were another focus of three different studies. Agullo and Bueno-Alastuey (2017), for instance, focused on the oral skills covered by these English language textbooks, such as listening, speaking, and pronunciation, Morales (2018) focused on listening skills, and Katawazai et al. (2019) focused on English sub-skills such as vocabulary, grammar, and pronunciation.

Additionally, visuals used in various English language textbooks were highlighted in two studies: Soori et al. (2011) evaluated a textbook's graphic representation, while Yazdanmehr and Soghi (2014) looked at a textbook's visualisation effects. Moreover, two studies centred on principles of Communicative Language Teaching employed by English language textbooks (Ahmad et al., 2019; Zeghdoud et al., 2019) and the knowledge dimension and cognitive processes that these textbooks capitalised on (Bakdash & Talebinejad, 2015; Amiri & Rezvani, 2021) respectively. Meanwhile, one study each focused on communicative competence elements (Caner & Celik, 2020), explicit instructions and implicit use of second language learning strategies (Bueno-Alastuey & Agullo, 2015), and themes related to UNESCO's global citizenship learning domains (Ait-Bouزيد, 2020).

Themes on textbook evaluation criteria in the selected studies

Nine themes on English language textbook evaluation criteria were identified after thoroughly analysing the selected studies in this review. These themes could be divided into three different groups: textbook evaluation criteria on the content (six), appearance (two), and supplementary materials (one) of the textbooks.

Content

1. Language learning activities/tasks

The most popular English language textbook evaluation criteria observed in the selected studies are those that evaluate language learning tasks and activities. Twenty-six studies focused on aspects related to these criteria (see Appendix 6). These criteria include task authenticity, whether they involve real tasks, types of tasks such as speaking and role-play tasks, and how textbook users will complete a task – in pairs, groups, or individually (Zeghdoud et al., 2019). In addition, Soori et al. (2011) focused on aspects related to communicative-based tasks in a textbook, while Caner and Celik (2020) looked at communicative strategies such as turn-taking, repetition, clarification requests, maintaining conversation, and the roles of participants, structures, examples, and instructions with regards to communicative tasks. On the other hand, Zara-ee and Hijazi (2018) highlighted the balance in language learning tasks and activities, while Hamidi et al. (2016) focused on the sequence and meaningfulness of such tasks and activities. Finally, Bueno-Alastuey and Agullo (2015) underlined the types of language learning strategies employed by these tasks.

2. Language skills

The second most popular criteria observed in the selected studies is the language skills covered by an English textbook. Twenty-five studies focused on aspects related to language skills (see Appendix 6), namely Reading, Writing, Speaking, and Listening, together with the sub-skills like Grammar, Pronunciation, and Vocabulary. The evaluation criteria also include aspects such as vocabulary repetition, the sequence of these skills according to the level, and whether these skills are presented in a meaningful way or not (Hamidi et al., 2016). In addition, Agullo and Bueno-Alastuey (2017) focused on how these different skills are integrated into English language textbooks and their authenticity for real-life purposes.

3. Subject / topic

The third most popular criteria observed in the selected studies are the subjects or topics of textbook content. Twenty studies focused on aspects related to these criteria (see Appendix 6). In general, aspects under these criteria evaluate the suitability, familiarity, and relation of the topics for the general content of the textbooks to students' daily life. In addition, Haghighi (2014), for example, looked at whether the topics were motivating and varied for students. Furthermore, Ait-Bouزيد (2020) underlined the themes of these topics concerning their connection to three global citizenship domains of learning which are cognitive, socio-emotional, and behavioral domains as framed by UNESCO's (2015) framework of global citizenship education.

4. Cultural

Eleven studies in this review highlighted cultural representation as one of their evaluation criteria (see Appendix 6). In general, aspects related to these criteria include the sources of the cultural content in the textbooks (target culture, local culture, international culture) and the categories of culture being dis-

cussed in them, whether they ‘big C’ culture (famous works of art, music, and literature, writers, artists, and musicians) or ‘little C’ culture (features of everyday life, beliefs, customs, behavior, and values).

Apart from that, Ashtiani (2013) looked at how these cultural contents were presented to students, whether implicitly or explicitly, with or without bias. The researchers also focused on how the cultural contents matched the students’ context and how interested the students were in learning about cultural or cross-cultural issues. Amerian and Tajabadi (2021), on the other hand, explored subcultures, taboos, and racial and gender stereotypes through their textbook evaluation criteria. They also evaluated attitudes behind cultural content, such as tolerance and empathy, challenging existing stereotypes, and arousing curiosity. Another researcher, Ariawan (2020), underlined social aspects of cultural representation in English language textbooks by examining elements such as social identity, social group, social interaction, social and political institutions, and social beliefs and behaviors.

5. Appropriateness for students

Eight of the selected studies in this review employed textbook evaluation criteria that looked at the appropriacy of English language textbooks for students (see Appendix 6). Soori et al. (2011), for instance, underlined the connection between the content of the textbooks and the objectives of the learners, their level of difficulty, their suitability for different types of learners, and whether the content provides ample opportunities for students to experience interactive learning while using the textbooks. On the other hand, Amiri and Rezvani (2021) attempted to determine the content suitability of the textbooks by using criteria guided by Bloom’s taxonomy (Bloom et al., 1956). These include aspects related to cognitive processes such as remembering and understanding, applying, analysing, evaluating, and creating knowledge dimensions such as factual, conceptual, procedural, and metacognitive, and how these aspects are represented in the textbooks to ensure that they are appropriate for the students.

6. Alignment with curriculum

The least popular English language textbook evaluation criteria observed in the selected studies is the alignment of the textbook with the goals and objectives of a country’s language programme and English language curriculum. Two studies focused on aspects related to these criteria (Alharbi, 2015; Marzban & Zokaeieh, 2017). The aspects include the systematic organisation of these objectives in English textbooks and whether these textbooks allow for different learning styles of their users.

Appearance

The appearance of textbooks has also become another focus in evaluating English language textbooks. Two criteria were determined to be among the most popular for evaluating these textbooks: the textbooks’ layout and design and practicality.

1. Layout and design

Eighteen studies in this review focused on criteria related to the layout and design of the textbooks (see Appendix 7). Among others, aspects related to these criteria tried to establish whether the textbooks’ layout and design, in general, are clear and well organised or not. Not only that, Hamidi et al. (2016), for instance, evaluated the clarity of the images, pictures, illustrations, and instructions. Likewise, Yazd-anmehr and Soghi’s (2014) evaluation criteria highlighted typography, visual arts, page layout, cover design, and physical makeup.

2. Practical considerations

Sixteen studies in our review focused on criteria concerning the practical considerations in using the textbooks (see Appendix 7), including the quality of instructions in the textbooks and their outline of clear learning objectives (Zara-ee & Hijazi, 2018), the textbooks' overall quality and durability (Soori et al., 2011), and their size (Hamidi et al., 2016).

Supplementary materials

It is not uncommon for publishers of textbooks to include supplementary materials with their textbooks. For teachers, these could come in the form of a teacher's edition, which contains the answers to the activities in the textbook, slides, test banks, and lecture notes. For students, these could include a workbook, CD, cassette, or a website or mobile application that hosts homework, tests, quizzes, and interactive audio-visual materials. Six studies in this review incorporated evaluation criteria on such materials (see Appendix 7). However, these criteria were only used in the studies selected in this review to indicate the inclusion of such materials to accompany the English language textbooks.

Discussion

At least two concerns about the criteria listed in textbook evaluation checklists reviewed by this study were uncovered. The first concern is how some of these checklists do not differentiate between analysis and evaluation; the checklists mix both procedures in a single checklist. Tomlinson (2013) argued that 'an evaluation is not the same as an analysis' (p. 22). According to Tomlinson, an evaluation makes the users of the materials its focus and creates judgments about their likely effects while being inevitably subjective, no matter how structured, criterion-referenced, and rigorous an evaluation is. On the other hand, an analysis that focuses on the materials and aims to provide an objective analysis 'asks questions about what the materials contain, what they aim to achieve, and what they ask learners to do' (Tomlinson, 1999, p. 10). Therefore, when such checklists combine analysis and evaluation, the questions in the checklists are more likely to be ineffective in evaluating an English language textbook since most of the analysis could 'be weighted disproportionately when combined with evaluation questions' (Tomlinson, 2013, p.22). This matter could jeopardise the validity of the checklists.

The second concern is about the validity of the criteria in the checklists. A valid instrument and criteria measure what they intend to measure. However, Nimehchisalem and Mukundan (2015) highlighted that some checklists are not tested for validity since they were developed hastily (Mukundan & Ahour (2010). In establishing valid and invalid criteria in checklists, the criteria, for instance, should 'encourage the analysis of materials based on the target situation of use so that they can be matched and evaluated for selection' (Mukundan & Ahour, 2010, p. 348). They then outlined three main features that textbook evaluation checklist developers should keep in mind when developing a checklist: clarity, conciseness, and flexibility.

Additionally, Tomlinson and Matsuhara (2004, cited by Tomlinson, 2013) both proposed five considerations in determining the validity of criteria for a textbook evaluation checklist, including ensuring (1) each question is an evaluation question, (2) each question only asks one question, (3) each question is answerable, (4) each question is free of dogma, and finally (5) each question will be interpreted in the same way by different evaluators.

Nevertheless, it must be noted that this review is not focused on identifying questions for evaluation and analysis or differentiating the validity of the criteria in the selected evaluation checklists. This review is intended to list the common criteria in these checklists and suggest additional evaluation criteria that can be included in future checklists.

Meanwhile, four additional evaluation criteria could be included for future English language textbook evaluation checklists: these are the relation of an English textbook to the goals and objectives of a country's language programme and curriculum, the continuity of language skills in an English textbook between different schooling years, a textbook's ability to support self-learning, and its technological integration.

First, more attention needs to be given to textbook evaluation criteria regarding the relationship between an English language textbook and the goals and objectives of a country's language programme and curriculum. There could be a detrimental effect on the effectiveness of the textbook to the users if it is not intended for a local audience. In Malaysia, for instance, where imported English textbooks are being used to teach English in its schools, it was revealed that almost 70 percent of English teachers believed that these textbooks have failed to support the goals and objectives of the country's language programme and curriculum (Din & Yamat, 2020). Although the books are aligned with the Common European Framework of Reference for Languages (CEFR), which the country has recently adopted for English language teaching, there are issues in its alignment with the goals and objectives of the country's English language teaching programme. Perhaps, instead of just including a general analysis question such as 'Does this textbook support the goals and objectives of the country's language programme and curriculum?', such criteria could be refined and expanded to include evaluation questions on how this relationship could affect the textbook's users.

Second, there is a need to include evaluation criteria on the continuity of language skills in English language textbooks from one school year to the next. In Malaysia, for instance, imported English language textbooks are published by different publishers each school year. This raises concerns about how each language skill tallies from one different publisher to another publisher, despite their claim to have followed the CEFR standards. This continuity aspect should be included as a criterion in textbook evaluation exercises, as the CEFR dictates that the 'process of language learning is continuous' (CoE, 2001, p. 17).

Third, as in-person learning in schools was halted, and online learning has become more prominent in the post-COVID-19 era, it is believed that a printed textbook could provide continuity for teaching and learning and equal access to knowledge and education, especially for those who come from lower to middle-income families (Chabbott & Sinclair, 2020; Shak et al., 2022) and those with device and connection issues (Bell et al., 2020; Buchbinder, 2020; Shak et al., 2022). This is due, among other considerations, to a textbook's ability to support self-directed learning (Cunningsworth, 1995; Rubdy, 2013; Ahmed, Yaqoob, & Yaqoob, 2015), where learning can take place beyond the classroom and at home (Jusuf, 2018). Therefore, there is a need for a textbook evaluation checklist to include criteria related to a textbook's capacity to support self-study during the pandemic era and beyond.

Still, concerning the learning situation in the post-pandemic world, a renewed spotlight is needed on the role of a textbook's supplementary materials to support online learning at home. Additional textbook evaluation criteria are essential to evaluate these materials. For instance, since learning at home is mainly assisted by caretakers, evaluation questions such as 'How likely will the teacher's edition be able to assist caretakers at home?' can be included. This is because these caretakers might not have adequate pedagogical knowledge to teach and help their children at home. What is more, in the context of English language learning, most of these caretakers are not fluent speakers of the English language. This could present another challenge for them to understand instructions in both the textbook and the teacher's guide.

Fourth, additional textbook evaluation criteria are recommended to evaluate technological integration with printed textbooks. Nowadays, it is not uncommon for a printed textbook to be accompanied by a companion website or mobile application to enhance user learning experiences. It is also not un-

sual for a textbook to provide links within its pages for students to access audio-visual materials or to participate in interactive learning activities beyond the book. Therefore, such technological integrations must be scrutinised through evaluation questions such as ‘Is the mobile application likely to enhance learning among the users?’ or ‘Is the mobile application likely to keep the users amused?’ This matter is particularly vital in the era of online learning, which depends almost entirely on technology for its success. The plethora of materials on the internet with dubious credibility could potentially harm or mislead our students.

Conclusion

This review examined the most common themes of textbook evaluation criteria used to evaluate English language textbooks from 2011 to 2021. Nine common themes were captured under three distinct components of textbook evaluation. In addition, four additional textbook evaluation criteria were also proposed to address the current English language learning situation.

Undeniably, these textbooks were published before the COVID-19 pandemic and were not intended to be used during this unprecedented period. They also might not rise to the challenges of helping learners in ensuring continuous learning during this pandemic (Shak et al., 2021a), although they are deemed to be one of the most convenient (Millar & Schrier, 2015) and cheapest ways (Ur, 1996) to access learning since they can be easily used without a device or internet connection. Not only that, but the English language textbook evaluation studies in this review were completed outside the context of the COVID-19 pandemic. Because of this, it is imperative to look again at how English language textbook evaluation practices and criteria should evolve in line with how our perceptions of the roles of English language textbooks have been altered in terms of teaching and learning in the post-pandemic world. This is essential to ensure that these textbooks can serve their fullest potential.

Biographical notes

Mohamad Syafiq Ya Shak

is a member of the Academy of Language Studies, Universiti Teknologi MARA (UiTM) Perak Branch, Malaysia. His research interests include material development and evaluation, digital learning, corpus-based studies, and media. He holds a Master of Education (TESL) and is currently pursuing his Ph.D. in Education (TESL), examining CEFR-aligned textbooks in Malaysian schools. Address: Universiti Teknologi Mara, Perak, Malaysia. Email: moham381@uitm.edu.my

Intan Safinas Mohd Arif Albakri

(corresponding author) is an associate professor lecturer from the Faculty of Languages and Communication, University Pendidikan Sultan Idris (UPSI), Tanjung Malim, Perak, Malaysia. Her areas of expertise are materials and methods in teaching the English Language. She is currently leading research on mentoring pre-service and beginning teachers and a member of research on Developing a Quality Teacher Education Framework under the Ministry of Education Malaysia. She is also the Editor-in-chief for the Asian Journal of English Language and Pedagogy. Address: Faculty of Languages and Communication, Universiti Pendidikan Sultan Idris, Perak, Malaysia. Email: safinas@upsi.edu.my

Siti Shuhaida Shukor

is a senior lecturer from the Faculty of Languages and Communication, University Pendidikan Sultan Idris (UPSI), Tanjung Malim, Perak, Malaysia. She specializes in Blended learning, Cultural-historical activity theory, social media, technology in Education. Address: Faculty of Languages and Communication, Universiti Pendidikan Sultan Idris, Perak, Malaysia. Email: sshuhaida@fbk.upsi.edu.my

Mohd Haniff Mohd Tahir

is a senior Lecturer and internationally recognized researcher at the Faculty of Languages and Communication, Universiti Pendidikan Sultan Idris (UPSI), Malaysia. His focus area includes materials development and evaluation, interactive language teaching curriculum design, digital learning, and quantitative research related to ESL classrooms. Currently, he is exploring emerging ideas and concepts for vocabulary learning. Address: Faculty of Languages and Communication, Universiti Pendidikan Sultan Idris, Perak, Malaysia. Email: haniff.tahir@upsi.edu.my

References

- Abbasian, R., & Biriya, R. (2017). English Language textbooks in EFL education: Do they improve students' national, international, and target culture familiarity? *Khazar Journal of Humanities and Social Sciences*, 20(2), 49–65. <https://doi.org/10.5782/2223-2621.2017.20.2.49>
- Abdelwahab, M.M. (2013). Developing an English Language textbook evaluative checklist. *IOSR Journal of Research & Method in Education*, 1(3), 55–70.
- Abd Rashid, R.S.A. & Engku Ibrahim, E.H. (2018). English language textbooks and portrayal of culture: A content analysis. *MATEC Web of Conferences*, 150, 1–12. <https://doi.org/10.1051/mateconf/201815005076>
- Agullo, G.L. & Bueno-Alastuey, M.C. (2017). Analysis of oral skills development in the Most used English language textbooks in the second year of the Baccalaureate in Spain. *Porta Linguarum*, 27, 107–121.
- Ahmad, M., Ismail, M.K.A., & Saba, N. (2019). Evaluating an English language textbook: A study on 'English-2' in the light of Communicative Language Teaching principles. *Revista Dilemas Contemporáneos: Educación, Política y Valores*, VI, 1–38.
- Ahmed, M., Yaqoob, H., & Yaqoob, M. (2015). Evaluation of listening skill of ELT textbook at secondary school level. *Advances in Language and Literary Studies*, 6(3), 225–229.
- Ahour, T., Towhidiyan, B., & Saeidi, M. (2014). The evaluation of 'English textbook 2' taught in Iranian high schools from teachers' perspectives. *English Language Teaching*, 7(3), 150–158. <https://doi.org/10.5539/elt.v7n3p150>
- Ait-Bouazid, H. (2020). Exploring global citizenship as a cross-curricular theme in Moroccan ELT textbooks. *Eurasian Journal of Applied Linguistics*, 6(2), 229–242. <https://doi.org/10.32601/ejal.775801>
- Akef, K. & Moosavi, Z. (2016). Iranian EFL teachers' and students' textbook evaluation. *The Iranian EFL Journal*, 10(6). <https://doi.org/10.13140/RG.2.1.5119.5763>
- Alharbi, A. (2015). A descriptive-evaluative study of a Saudi EFL textbook series. *Cogent Education*, 2(1). <https://doi.org/10.1080/2331186X.2015.1079946>
- Al-Sowat, H. (2012). An evaluation of the English language textbook 'Say it in English' for first-year intermediate grade in Saudi Arabia. *Journal of Studies in Curriculum and Supervision*, 3(2), 332–413.
- Amerian, M., & Tajabadi, A. (2020). The role of culture in foreign language teaching textbooks: An evaluation of New Headway series from an intercultural perspective. *Intercultural Education*, 31(6), 623–644. <https://doi.org/10.1080/14675986.2020.1747291>
- Amiri, A., & Rezvani, R. (2021). A tale of three official English textbooks: An evaluation of their horizontal and vertical alignments. *Language Related Research*, 12(3), 51–79. <https://doi.org/10.29252/LRR.12.3.3>
- Ansary, H. & Babaii, E. (2002). Universal characteristics of EFL/ESL textbooks: A step towards systematic textbook evaluation. *The Internet TESL Journal*, 8(2), available online at: <http://iteslj.org/Ansary-textbooks/>
- Ariawan, S. (2020). Investigating cultural dimensions in EFL textbook by using Byram Checklist. *Register Journal*, 13(1), 123–152. <https://doi.org/10.18326/rgt.v13i1.123-152>

- Ashtiani, P.A. (2013). American Headway Starter: A textbook evaluation. *Modern Journal of Language Teaching Methods*, 3(2), 30–36.
- Azadsarv, M., Taghvaei, M., & Zamani, S. (2015). Cultural exophoric references in English textbooks used in four common English textbooks used at Iranian English Language Institutes: A comparative discursive study of the degree of referential burden. *Journal of Research in Applied Linguistics*, 5(Special Issue), 66–75.
<https://doi.org/10.22055/rales.2015.11270>
- Azarfam, A.A.Y., & Noordin, N. (2014). Evaluating an English textbook for application in Iranian EFL academic context. *Theory and Practice in Language Studies*, 4(3), 618–623.
<https://doi.org/10.4304/tpls.4.3.618-623>
- Baktash, N. & Talebinejad, M.R. (2015). Evaluation of the new Iranian high school series books based on Bloom's revised taxonomy: Prospect One in focus. *Modern Journal of Language Teaching Methods*, 5(3), 439–445.
- Bell, S., Cardoso, M., Giraldo, J.P., El Makkouk, N., Nasir, B., Mizunoya S., & Dreesen, T. (2020). Can broadcast media foster equitable learning amid the COVID-19 pandemic? *UNICEF Connect*. Retrieved from <https://blogs.unicef.org/evidence-for-action/can-broadcast-media-foster-equitable-learning-amid-the-covid-19-pandemic/>
- Bloom, B.S., Engelhart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. (1956). *Taxonomy of educational objectives: The classification of educational goals*. Handbook I: Cognitive domain. Longman.
- Brown, D.H. (2001). *Teaching by principles: An interactive approach to language pedagogy*. 2. ed. Longman.
- Braun, V. & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sports Exercise Health*, 11(4), 589–597.
<https://doi.org/10.1080/2159676X.2019.1628806>
- Buchbinder, N. (2020). *Digital capacities and distance education in times of coronavirus: Insights from Latin America*. World Education Blog. UNESCO. Retrieved from <https://gemreportunesco.wordpress.com/2020/05/12/digital-capacities-and-distance-education-in-times-of-coronavirus-insights-from-Latin-America/>
- Bueno-Alastuey, M.C., & Agulló, G.L. (2015). Explicit instruction and implicit use of L2 learning strategies in Higher Secondary EFL course books. *International Journal of English Studies*, 15(2), 17–39.
<https://doi.org/10.6018/ijes/2015/2/218581>
- Byram, M. (1993). *Language and culture: The need for integration*. In M. Byram (Ed.), Germany: Its representation in textbooks for teaching German in Great Britain (pp. 3–16). Diesterweg.
- Caner, M., & Çelik, A. (n.d.). The Communicative Competence elements in the foreign language textbooks: A descriptive case study on Turkish and English textbooks. *International Association for Research on Textbooks and Educational Media*, 12(1), 1–19
- Chabbott, C. & Sinclair, M. (2020). SDG 4 and the COVID-19 emergency: Textbooks, tutoring, and teachers. *Prospects*, 49, 51–57.
<https://doi.org/10.1007/s11125-020-09485-y>
- CoE (Council of Europe). (2001). *Common European Framework of Reference for languages: Learning, teaching, assessment*. Cambridge University Press.
- Cunningsworth, A. (1995). *Choosing your coursebook*. Heinemann English Language Teaching.
- Day, R.R., & Park, J.-S. (2005). Developing reading comprehension questions. *Reading in a Foreign Language*, 17(1), 60–73.
- Daoud, A., & Celce-Murcia, M. (1979). *Selecting and evaluating a textbook*. In M. Celce-Murcia & L. McIntosh (Eds.), Teaching English as a second or foreign language (pp. 302–307). Newbury House Publishers.
- Demir, Y. & Ertas, A. A. (2014). A suggested eclectic checklist for ELT coursebook evaluation. *The Reading Matrix*, 14(2), 243–252.
- Dos Santos, L.M. (2020). The application of a textbook evaluation checklist: A research study of English as a foreign language textbook. *Universal Journal of Educational Research*, 8(9), 3864–3872.
<https://doi.org/10.13189/ujer.2020.080910>
- Din, F.F.M & Yamat, H. (2020). Challenges faced by teachers in using the Supermind textbook. *International Journal of Academic Research in Progressive Education and Development*, 9(1), 122–133
- Ghezlou, M., Moghtadi, L., & Koosha, M. (2016). EFL teachers' and students' perspectives on Total English Upper Intermediate: An analytical evaluation. *Modern Journal of Language Teaching Methods*, 6(2), 101–116.
- Ghorbani, M.R. (2011). Quantification and graphic representation of EFL textbook evaluation results. *Theory and Practice in Language Studies*, 1(5), 511–520.

- Haghighi, S. (2014). Evaluation of 'First Certificate Gold Coursebook': Evidence from students of a private English Institute in Iran. *International Journal of Instruction*, 7(2), 149–164.
- Hamidi, H., Bagheri, M., Sarinavaee, M., & Seyyedpour, A. (2016). Evaluation of two general English textbooks: New Interchange 2 vs. Four Corners 3. *Journal of Language Teaching and Research*, 7(2), 345–351.
<https://doi.org/10.17507/jltr.0702.13>
- Jusuf, H. (2018). The models of checklist method in evaluating ELT textbooks. *Journal Al-Lisan*, 3(2), 17–35.
- Karimi, S., Kargar, A.A., & Behjat, F. (2015). An evaluation of commonly used general English textbooks of Iranian university contexts. *Modern Journal of Language Teaching Methods*, 5(3), 257–264.
- Katawazai, R., Haidari, M., & Sandaran, S. C. (2019). An evaluation of sub-skills (vocabulary, grammar, and pronunciation) in Grade 9 English textbook of Afghan secondary schools. *International Journal of Engineering and Advanced Technology*, 8(5), 1236–1241. <https://doi.org/10.35940/ijeat.E1176.0585C19>
- Kausar, G., Mushtaq, M., & Badshah, I. (2016). The evaluation of English language textbook taught at an intermediate level. Gomal University. *Journal of Research*, 4, 32–43.
- Keban, N. V. (2012). *Content analysis on English for kids Grade 3, a textbook used in elementary schools in Malang*. [Unpublished Master's Thesis]. Universiti Malang
- Khodabakhshi, M. (2014). Choose a proper EFL textbook: Evaluation of 'Skyline' Series. *Procedia - Social and Behavioral Sciences*, 98(1988), 959–967.
<https://doi.org/10.1016/j.sbspro.2014.03.505>
- Kiger, M.E. & Varpio, L. (2020). Thematic analysis of qualitative data. *Medical Teacher*, 42(8), 846–854.
<https://doi.org/10.1080/0142159X.2020.1755030>
- Larrea-Espinar, Á., & Raigón-Rodríguez, A. (2019). The development of culture in English foreign language textbooks: The case of English File. *Revista de Lenguas Para Fines Específicos*, 25(2), 114–133.
<https://doi.org/10.20420/rife.2019.384>
- Lee, K.Y. (2009). Treating culture: What 11 high school EFL conversation textbooks in South Korea do. *English Teaching: Practice and Critique* 8(1): 76–96.
- Litz, D.R.A. (2005). Textbook evaluation and ELT management: A South Korean case study. *Asian EFL Journal*, 48, 1–53.
- Majid, S.N.A., & Salam, A.R. (2021). A systematic review of augmented reality applications in language learning. *International Journal of Emerging Technologies in Learning (IJET)*, 16(10), 18–34.
<https://doi.org/10.3991/ijet.v16i10.17273>
- Maleki, A., Mollaei, F., & Khosravi, R. (2014). A content evaluation of Iranian pre-university ELT textbook. *Theory and Practice in Language Studies*, 4(5), 995–1000.
<https://doi.org/10.4304/tpls.4.5.995-1000>
- Marzban, A., & Zokaeieh, S. (2017). Local evaluation of an EFL textbook: 'Mosaic 1: Listening and Speaking.' *Theory and Practice in Language Studies*, 7(10), 874–882.
<https://doi.org/10.17507/tpls.0710.08>
- Mieky, J. (2005). EFL Evaluation Checklist. *The Reading Matrix*, 5(2), 1–9.
- Millar, M., & Schrier, T. (2015). Digital or printed textbooks: Which do students prefer and why? *Journal of Teaching in Travel & Tourism*, 15(2), 166–185.
<https://doi.org/10.1080/15313220.2015.1026474>
- Moher, D., Shamsee, L., Clarke, M., Ghersi, D., Liberati, A., Petticrew, M., Shekelle, P., & Stewart, L.A. (2015). Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. *Systematic Reviews*, 4, 1.
<https://doi.org/10.1186/2046-4053-4-1>
- Momand, M., Sugunabalan, A., Ibrahim, S.A.M., & Sandaran, S.C. (2019). An evaluation of primary school year 6 (KSSR) English textbook. *International Journal of Engineering and Advanced Technology*, 8(5), 1192–1199.
<https://doi.org/10.35940/ijeat.E1169.0585C19>
- Monazzah, Z., Behjat, F., & Zareian, A. (2016). General English university coursebook evaluation based on Litz's model and professors' perceptions of such books. *Modern Journal of Language Teaching Methods*, 6(2), 435–442
- Morales, B. C. (2018). Types of listening comprehension promoted in the Chilean EFL textbook Global English. *Inglés Global*, 20(1), 49–61

- Mukundan, J., & Ahour, T. (2010). *A review of textbook evaluation checklists across four decades (1970–2008)*. In B. Tomlinson, & H. Masuhara (Eds.), *Research for materials development in language learning: Evidence for best practice* (pp. 336–352). Continuum.
- Mukundan, J., Nimehchisalem V., & Hajimohammadi R. (2011). Developing an English language textbook evaluation checklist: A focus group study. *International Journal of Humanities and Social Science*, 1(12), 100–106
- Nation, I.S.P. & Macalister, J. (2010). *Language curriculum design*. Routledge.
- Nikou, F.R., Ramzi, S., & Tabatabaei, M.A. (2014). Contrasting two Iranian ESP textbooks written by TEFL experts and field specialists. *The Journal of Teaching English for Specific and Academic Purposes*, 2(1), 101–109
- Nimehchisalem, V. & Mukundan, J. (2015). Refinement of the English language teaching textbook evaluation checklist. *Pertanika Journal of Social Sciences & Humanities*, 23 (4), 761–780
- Nourmohammad-Nouri, N., Ghaemi, F., & Sedigh, R. (2015). 'English Result': The qualitative and quantitative analysis of course books as perceived by Iranian EFL learners and teachers. *Procedia – Social and Behavioral Sciences*, 192, 601–614. <https://doi.org/10.1016/j.sbspro.2015.06.106>
- Oxford, R. (1990). *Language learning strategies: What every teacher should know*. Heinle.
- Paige, R.M., Jorstad, H., Paulson, L., Klein, F., & Colby, J. (1999). *Culture learning in language education: A review of the literature. Culture as the core: Integrating culture into the language curriculum (CARLA Working Paper Series 15)*. Eds. R. Michael Paige, D. Lange, and Y. Yershova. Minneapolis: The Center for Advanced Research on Language Acquisition, University of Minnesota.
- Raigón-Rodríguez, A. (2018). Analysing cultural aspects in EFL textbooks: A skill-based analysis. *Journal of English Studies*, 16, 281–300. <https://doi.org/10.18172/jes.3478>
- Razmjoo, S. A. (2007). High schools or private institutes textbooks? Which fulfill communicative language teaching principles in the Iranian context? *Asian EFL Journal*, 9(4), 126–140.
- Rubdy, R. (2014). *Selection of materials*. In B. Tomlinson (Ed.), *Developing materials of language teaching: Chapters from the first edition* (2nd ed., pp. 37–57). Bloomsbury.
- Samsudin, S.F., Masenwat, N.Z., & Ab Llah, L. (2021). Systematic literature review on cultivating reading culture in academic library initiatives. *Journal of Academic Library Management*, 1(2), 5674, 2021. <https://doi.org/10.24191/aclim.v1i2.13>
- Shak, M.S.Y., Abdul Malik, N., Hasni, N.A. & Mohd Tahir, M.H. (2021a). Revisiting PULSE 2 textbook after three years: Have teachers changed their minds? *Malaysian Journal of Social Sciences and Humanities*, 6(5), 177–189. <https://doi.org/10.47405/mjssh.v6i5.777>
- Shak, M.S.Y., Albakri, I.S.M.A., Tahir, M.H.M., & Adam, M.H.M. (2021b). The use of imported CEFR-aligned English Language textbooks in Malaysian schools: Issues and concerns. *International Journal of Academic Research in Business and Social Sciences*, 11(9), 954–963
- Shak, M.S.Y., Tahir, M.H.M., & Ibrahim, N. (2022). Printed textbook and its potential in remote learning during the Covid-19 Crisis. *International Journal of Academic Research in Progressive Education and Development*, 11(1), 403–412.
- Shatery, H. & Azargoon, M. (2012). *Designing and developing a native checklist to evaluate general English course books in Iran and comparing it with other existing checklists in the world*. Conference on Language Learning and Teaching: An Interdisciplinary Approach, October 30–31, 2012. Retrieved from <https://www.sid.ir/paper/927171/en>
- Soori, A., Kafipour, R., & Soury, M. (2011). EFL textbook evaluation and graphic representation. *European Journal of Social Sciences*, 26(3), 481–493
- Tamilchelvan, S., and Rashid, R. A. (2017). Being a Muslim gay man: A systematic review. *TRAMES: A Journal of the Humanities & Social Sciences*. 21, 273–284 <https://doi.org/10.3176/tr.2017.3.05>
- Tomlinson, B. (1999). Developing criteria for materials evaluation, *IATEFL Issues*, 147, 10–13
- Tomlinson, B. (2013). *Materials evaluation*. In B. Tomlinson (Eds.), *Developing materials for language teaching* (pp. 21–48). Bloomsbury Publishing Plc.
- UNESCO. (2015). *Global citizenship education: Topics and learning objectives*. United Nations Educational, Scientific, and Cultural Organization
- Ur, P. (1996). *A course in language teaching: Practice and theory*. Cambridge University Press.
- Williams, D. (1983). Developing criteria for textbook evaluation. *ELT Journal*, 37, 251–255. <http://dx.doi.org/10.1093/elt/37.3.251>

Yazdanmehr, E. & Shoghi, S. (2014). Design and application of a 'textbook visual effects' evaluation checklist. *Theory and Practice in Language Studies*, 4(3), 473–480,

Zara-ee A. & Hejazi, S.Y. (2018). Iranian university learners' and teachers' views on adopted and locally developed English language teaching textbooks. *International Journal of Instruction*, 11(3), 291–308.

Zeghdoud, M., Al-Haq, F.A.A., & Al-Jamal, D. (2019). Communicative English language teaching principles in action: Moving from theory to examples. *Jordan Journal of Modern Languages and Literatures*, 11(2), 225–246.

Appendix 1

Sample textbook evaluation checklist by Mukundan et al. (2011)

I. GENERAL ATTRIBUTES					
A. The book in relation to syllabus and curriculum					
1. It matches to the specifications of the syllabus.	0	1	2	3	4
B. Methodology					
2. The activities can be fully exploited and embrace the various methodologies in ELT.	0	1	2	3	4
3. Activities can work well with methodologies in ELT.	0	1	2	3	4
C. Suitability to learners					
4. It is compatible with the age of the learners.	0	1	2	3	4
5. It is compatible with the needs of the learners.	0	1	2	3	4
6. It is compatible with the interests of the learners.	0	1	2	3	4
D. Physical and utilitarian attributes					
7. Its layout is attractive.	0	1	2	3	4
8. It indicates the efficient use of text and visuals.	0	1	2	3	4
9. It is durable.	0	1	2	3	4
10. It is cost-effective.	0	1	2	3	4
E. Efficient outlay of supplementary materials					
11. The book is supported efficiently by essentials like au-dio materials.	0	1	2	3	4
II. LEARNING-TEACHING CONTENT					
A. General					
1. Most of the tasks in the book are interesting.	0	1	2	3	4
2. Tasks move from simple to complex.	0	1	2	3	4
3. Task objectives are achievable.	0	1	2	3	4
4. Cultural sensitivities have been considered.	0	1	2	3	4
5. The language in the textbook is natural and real.	0	1	2	3	4
6. The situations created in the dialogues sound natural and real.	0	1	2	3	4
B. Listening					
7. The book has appropriate listening tasks with well-defined goals.	0	1	2	3	4
8. Tasks are efficiently graded according to complexity.	0	1	2	3	4
9. Tasks are authentic or close to real language situations.	0	1	2	3	4
C. Speaking					
10. Activities are developed to initiate meaningful communication.	0	1	2	3	4
11. Activities are balanced between individual response, pair work and group work.	0	1	2	3	4
D. Reading					
12. Texts are graded.	0	1	2	3	4
13. Texts are interesting.	0	1	2	3	4
E. Writing					
14. Tasks have achievable goals and take into consideration learner capabilities.	0	1	2	3	4
15. Tasks are interesting.	0	1	2	3	4

F. Vocabulary					
16. The load (number of new words in each lesson) is ap-propriate to the level.	0	1	2	3	4
17. There is a good distribution (simple to complex) of vo-cabulary load across chapters and the whole book.	0	1	2	3	4
18. Words are efficiently repeated and recycled across the book.	0	1	2	3	4
G. Grammar					
19. The spread of grammar is achievable.	0	1	2	3	4
20. The grammar is contextualized.	0	1	2	3	4
21. Examples are interesting.	0	1	2	3	4
22. Grammar is introduced explicitly and reworked inci-dentally throughout the book.	0	1	2	3	4
H. Pronunciation					
23. It is contextualized.	0	1	2	3	4
24. It is learner-friendly with no complex charts.	0	1	2	3	4
I. Exercises					
25. They are learner-friendly.	0	1	2	3	4
26. They are adequate.	0	1	2	3	4
27. They help students who are under/over-achievers.	0	1	2	3	4

Appendix 2

List of studies by country

Country	Studies
Iran	Soori et al., 2011; Ashtiani, 2013; Khodabakshi, 2014; Maleki et al., 2014; Ahour et al., 2014; Haghighi, 2014; Nikou et al., 2014; Yazdanmehr & Soghi, 2014; Nourmohammad-Nouri et al., 2015; Azadsarv et al., 2015; Ghezlou, et al., 2015; Baktash & Talebinejad, 2015; Hamidi et al., 2016; Monazzah et al., 2016; Karimi, et al., 2016; Marzban & Zokaeieh, 2017; Abbasian & Biria, 2017; Zara-ee & Hijazi, 2018; Azarfam & Noordin, 2018; Amerian & Tajab-adi, 2020; Amiri & Rezvani, 2021
Spain	Bueno-Alastuey & Agullo, 2015; Agullo & Bueno-Alastuey, 2017; Raigon-Rodriguez, 2018; Larrea-Espinar & Raigon-Rodriguez, 2019
Malaysia	Abd Rashid & Engku Ibrahim, 2017; Momand et al., 2019
Afghanistan	Katawzai et al., 2019
Algeria	Zeghdoud et al., 2019
Chile	Morales, 2018
Indonesia	Ariawan, 2020
Morocco	Ait-Bouzid, 2020
Pakistan	Ahmad et al., 2019
Saudi Arabia	Alharbi, 2015
South Korea	Dos Santos, 2020
Turkey	Caner & Celik, 2020

Appendix 3

Sources of checklists

Source of checklist	Studies
Adapted – Litz (2005)	Ahour et al., 2014; Khodabakshi, 2014; Ghezlou et al., 2015; Nourmohammad-Nouri et al., 2015; Monazzah et al., 2016; Ahmad et al., 2019
Adapted – Byram (1993)	Abbasian & Biria, 2017; Abd Rashid & Engku Ibrahim, 2017; Larrea-Espinar & Raigon-Rodriguez, 2019; Ariawan, 2020
Adapted – Bloom (2001)	Baktash and Talebinejad, 2015; Amiri and Rezvani, 2021
Adapted – Miekley (2005)	Maleki et al., 2014; Karimi et al., 2016
Adapted – Akef & Moosavi (2014),	Momand et al., 2019
Adapted – Al-sowat (2012)	Marzban & Zokaeieh, 2017
Adapted – Ansary and Babaii (2002)	Ashtiani, 2013
Adapted – Brown (2001)	Zeghdoud et al., 2019
Adapted – Daoud and Celce-Murcia (1979)	Hamidi et al., 2016
Adapted – Day and Park (2005)	Morales, 2018
Adapted – Demir and Ertas (2014)	Katawazai et al., 2019
Adapted – Ghorbani (2011)	Azarfam and Noordin, 2018
Adapted – Mukundan et al. (2011), Razmjoo (2007), Nation and Mcalister (2010)	Zara-ee and Hijazi, 2018
Adapted – Oxford (1990)	Bueno-Alastuey and Agullo, 2015
Adapted – Page et al. (1999) and Lee (2009)	Raigon-Rodriguez, 2018
Adapted – Shatery and Azarsoon (2012)	Azadsarv et al., 2015
Adapted – Williams (1983) and Keban (2012)	Alharbi, 2015
Own checklist	Soori et al., 2011; Haghighi, 2014; Nikou et al., 2014; Yazdanmehr and Soghi, 2014; Agullo and Bueno-Alastuey, 2017; Ait-Bouزيد, 2020; Amerian and Tajabadi, 2020; Caner and Celik, 2020; Dos Santos, 2020

Appendix 4:

Participants in the textbook evaluation exercises

Participants	Studies
Researchers	Soori et al., 2011; Ashtiani, 2013; Baktash & Talebinejad, 2015; Bueno-Alastuey & Agullo, 2015; Abbasian & Biria, 2017; Abd Ra-shid & Engku Ibrahim, 2017; Agullo & Bueno-Alastuey, 2017; Azarfam & Noordin, 2018; Raigon-Rodriguez, 2018; Ahmad et al., 2019; Zeghdoud et al., 2019; Larrea-Espinar & Raigon-Rodriguez, 2019; Katawazai et al., 2019; Caner & Celik, 2020; Ait-Bouزيد, 2020; Ariawan, 2020; Amiri & Rezvani, 2021
Teachers and students	Maleki et al., 2014; Haghighi, 2014; Azadsarv et al., 2015; Nourmo-hammad-Nouri et al., 2015; Ghezlou et al., 2015; Zara-ee & Hijazi, 2018; Dos Santos, 2020
Teachers only	Khodabakshi, 2014; Ahour et al., 2014; Alharbi, 2015; Hamidi et al., 2016; Karimi et al., 2016; Amerian & Tajabadi, 2020

Student only	Yazdanmehr & Soghi, 2014
Teachers and post-graduate students	Momand et al., 2019
Teacher trainees	Marzban & Zokaeieh, 2017
English teaching ex-perts	Nikou et al., 2014; Monazzah et al., 2016; Morales, 2018

Appendix 5:

Focus of the selected studies

Focus	Studies
Evaluating general features of English Language textbooks (e.g. their physical appearance, content, language skills, language tasks, exercises, and activities, practicality, topics, and supplementary materials).	Ashtiani, 2013; Khodabakshi, 2014; Maleki et al., 2014; Ahour et al., 2014; Haghighi, 2014; Nikou et al., 2014; Nourmohammad-Nouri et al., 2015; Alharbi, 2015; Ghezlou et al., 2015; Hamidi et al., 2016; Monazzah et al., 2016; Karimi et al., 2016; Marzban & Zokaeieh, 2017; Zara-ee & Hijazi et al., 2018; Azarfam & Noordin, 2018; Momand et al., 2019; Dos Santos, 2020

Appendix 6:

Textbook evaluation criteria: content

Textbook evaluation criteria	Studies
Language learning tasks and activities in an English language textbook	Khodabakshi, 2014; Zeghdoud et al., 2019; Zara-ee & Hijazi, 2018; Soori et al., 2011; Nourmohammad-Nouri et al., 2015; Morales, 2018; Momand et al., 2019; Marzban & Zokaeieh, 2017; Maleki et al., 2014; Hamidi et al., 2016; Dos Santos, 2020; Caner & Celik, 2020; Bueno-Alastuey & Agullo, 2015; Azarfam & Noordin, 2018; Ashtiani, 2013; Alharbi, 2015; Ahour et al., 2014; Ahmad et al., 2019; Amiri & Rezvani, 2021; Haghighi, 2014; Nikou et al., 2014; Monazzah et al., 2016; Karimi et al., 2016; Ghezlou et al., 2015; Baktash & Talebinejad, 2015; Agullo & Bueno-Alastuey, 2017
Language skills covered by an English textbook.	Khodabakshi, 2014; Zeghdoud et al., 2019; Zara-ee & Hijazi, 2018; Soori et al., 2011; Nourmohammad-Nouri et al., 2015; Morales, 2018; Momand et al., 2019; Marzban & Zokaeieh, 2017; Maleki et al., 2014; Katawazai et al., 2019; Hamidi et al., 2016; Dos Santos, 2020; Caner & Celik, 2020; Bueno-Alastuey & Agullo, 2015; Azarfam & Noordin, 2018; Ashtiani, 2013; Alharbi, 2015; Ahour et al., 2014; Ahmad et al., 2019; Haghighi, 2014; Nikou et al., 2014; Monazzah et al., 2016; Karimi et al., 2016; Ghezlou et al., 2015; Agullo & Bueno-Alastuey, 2017
Subjects or topics of content covered by an English textbook	Khodabakshi, 2014; Zara-ee & Hijazi, 2018; Nourmohammad-Nouri et al., 2015; Momand et al., 2019; Abd Rashid & Engku Ibrahim, 2017; Marzban & Zokaeieh, 2017; Maleki et al., 2014; Larrea-Espinar & Raigon-Rodriguez, 2019; Azarfam, Noordin, 2018; Ashtiani, 2013; Alharbi, 2015; Amerian & Tajabadi, 2020; Ait-Bouzid, 2020; Ahour et al., 2014; Abbasian & Biria, 2017; Ariawan, 2020; Haghighi, 2014; Monazzah et al., 2016; Karimi et al., 2016; Ghezlou et al. 2015.
Cultural content	Momand et al., 2019; Abd Rashid & Engku Ibrahim, 2017; Marzban and Zokaeieh, 2017; Larrea-Espinar & Raigon-Rodriguez, 2019; Ashtiani, 2013; Azadsarv et al., 2015; Alharbi, 2015; Amerian & Tajabadi, 2020; Abbasian, Biria, 2017; Ariawan, 2020; Karimi et al., 2016
Appropriacy of an English textbook for its users	Soori et al., 2011; Maleki et al., 2014; Hamidi et al., 2016; Dos Santos, 2020; Azarfam & Noordin, 2018; Ashtiani, 2013; Alharbi, 2015; Karimi et al., 2016; Amiri & Rezvani, 2021
Fitness of the textbook with the goals and objectives of a country's language programme and English language curriculum	Alharbi, 2015; Marzban & Zokaeieh, 2017

Appendix 7:

Textbook evaluation criteria: other aspects

Textbook evaluation criteria	Studies
Layout & design	Khodabakshi, 2014; Zara-ee & Hijazi, 2018; Soori et al., 2011; Nourmohammad-Nouri et al., 2015; Momand et al., 2019; Marzban & Zokaeieh, 2017; Hamidi et al., 2016; Dos Santos, 2020; Azarfam & Noordin, 2018; Ashtiani, 2013; Alharbi, 2015; Ahour et al., 2014; Haghighi, 2014; Nikou et al., 2014; Yazdanmehr & Soghi, 2014; Monazzah et al., 2016; Karimi et al., 2016; Ghezlou et al., 2015
Practical considerations	Khodabakshi, 2014; Zara-ee & Hijazi, 2018; Soori et al., 2011; Nourmohammad-Nouri et al., 2015; Momand et al., 2019; Hamidi et al., 2016; Dos Santos, 2020; Azarfam & Noordin, 2018; Ashtiani, 2013; Alharbi, 2015; Ahour et al., 2014; Haghighi, 2014; Nikou et al., 2014; Monazzah et al., 2016; Karimi et al., 2016; Ghezlou et al., 2015; Zeghdoud et al., 2019
Supplementary materials	Soori et al., 2011; Momand et al., 2019; Marzban & Zokaeieh, 2017; Azarfam & Noordin, 2018; Ashtiani, 2013; Alharbi, 2015