

**Evaluating ELT Materials with Specific Reference to Colleges of
Applied Sciences General Foundation Programme in Oman:
Towards a Viable checklist**

MUNA MAHAD DABALLAN KASHOOB

Doctor of Philosophy

University of York

Education

January 2018

Abstract

The main question and purpose of this study was to pinpoint a sustainable method on developing an evaluation checklist for teaching materials in the English language programmes. The answer of the main question is rejoined through three sub-questions about the sources for the checklist, identifying the design guidelines and offering a validation method for the developed checklist. Design-based research methodology was utilized through three main phases: analysis and exploration, design and construction and evaluation and reflection where the last phase comprised cycles of formative review of the developed checklist. The participants were purposefully sampled from the six Colleges of Applied Sciences in Oman and other higher education institutions. A report is written after each cycle of formative review (expert review, one –to-one, small group, and field testing) with the recommended changes which led to four revisions and redesigns of the checklist prototype. The results of this study were fourfold. First, a conceptual framework was designed that can be used to develop checklists for the evaluation of teaching materials in the English language programmes. Second, a verified checklist is developed that can be used as both an evaluation & selection instrument as well as a professional development tool. Third, formative review is perceived to be a powerful validation tool for reviewing the developed teaching materials evaluation checklists. Finally, guidelines on how to develop teaching materials evaluation checklists are yielded through the different phases of this study. The use of design-based research facilitated the design and the assessment of the checklist which indicates the necessity of such methodology in the complicated educational milieus with its focus on research, design and setting in unison. Besides the practical results, findings comprised adding new visualization of sources, content and use of teaching materials evaluation checklists. It is concluded that teaching materials evaluation is the main contributor to the students' progression, the practitioners' professional development and the success of the English language programmes as a whole.

Table of Contents

Abstract	ii
Table of Contents	iii
List of Tables.....	viii
List of Figures	ix
Acknowledgement.....	x
Declaration	xii
Dedication:	xiii
Chapter 1 Introduction to the Study	1
1.1. Study Rationale	1
1.1.1 Teaching materials in English language programmes.....	1
1.1.2 Selection of coursebooks.....	3
1.1.3 Materials Evaluation and Language Programmes.....	4
1.2 The Need for a Feasible Evaluation	9
1.3 Procedures of Developing Teaching Materials Evaluation Instruments	16
1.4 Overview of the Context.....	18
1.5 Education Policies, problems and Reforms	20
1.6 Colleges of Applied Sciences in Oman	23
1.6.1 The English Language Programmes: an Overview.....	24
1.6.2 The Foundation Programmes in the Omani Higher Education Institutions ...	25
1.6.3 The Foundation Programmes National Standard	26
1.6.4 Foundation Programmes Council Audits	28
1.6.5 Foundation Programmes Structure in the Colleges of Applied Sciences.....	30
1.7 The Main Phases of the Study	33
1.8 Research Questions	36
Chapter 2 English Language Teaching Materials Evaluation in the English Language Programmes.....	37
2.1 Introduction	37
2.2 An Overview of Evaluation Theories and Models	38
2.3 ELT Materials Evaluation clarified	42
2.4 Teaching Materials Definitions	44
2.5 Materials Design and Development	45
2.6 The Role Of Course Books in English Language Programmes	50
2.7 How and what to evaluate?.....	52

2.8 Evaluation and Teaching Materials in English Language programmes	56
2.9 Materials Evaluation and Types	61
2.9.1 Macro and Micro Evaluation	62
2.9.2 Pre-use, in-use and post-use Evaluation.....	63
2.9.3 External and Internal Evaluation.....	64
2.9.4 Formative and Summative Evaluation	64
2.10 Practical Methods for Teaching Materials Evaluation	65
2.10.1 The impressionistic method	66
2.10.2 The in-depth method	66
2.10.3 The checklist method	66
2.11 The Initial Concept of the Checklist and its Sources	74
2.12 Summary.....	78
Chapter 3: Research Design and Methodology.....	79
3.1 Introduction	79
3.2 A Conceptual Framework for the Evaluation Checklist Development	80
3.3 DBR within Common Research Paradigms	81
3.3.1. Introduction.....	81
3.3.2 DBR in a Critical Realism Paradigm	82
3.3.3. DBR in a Pragmatism Paradigm	83
3.4 Design-based Research and other Types of Research	85
3.4.1 DBR and Design	85
3.4.2 DBR and Action Research	87
3.4.3 DBR and Evaluation	88
3.4.4 Summary	89
3.5 Rationale for Design-Based Research	90
3.6 Design-Based Research definitions and models.....	92
3.6.1 Definitions.....	92
3.6.2 Models of Design-Based Research	94
3.7 The Organization of the Three Core Phases in This Study	97
Chapter 4: The Development of the Teaching Materials Evaluation Checklist.....	99
4.1 Introduction	99
4.2 The Conceptual Framework	99
4.2.1 Developing the Conceptual Framework.....	100
4.2.2 Testing the Conceptual Framework by Subject Matter Experts.....	102
4.2.3 Restructuring the Framework.....	104

4.3 The Main Constructs of the Checklist: Research	106
4.3.1 Second Language Acquisition Principles.....	106
4.3.2 Teaching Principles.....	110
4.3.3 The ELT Curriculum Design	114
4.4 Main Constructs of the Checklist: Practical Needs	119
4.4.1 Teachers, Students and Institutional Needs.....	119
4.4.2 The Instruments Used for Needs Assessment	121
4.4.3 Needs Analysis Instruments: Students Brainwriting Sessions	124
4.4.4 Needs Analysis- Considering the Students Data.....	125
4.4.5 Needs Analysis Instruments: Teachers Brainwriting Sessions	127
4.4.6 Needs Analysis- Considering Teachers Data	128
4.4.7 Institutional Needs Short Survey.....	130
4.4.8 Summary	131
Chapter 5 The Checklist Review.....	134
5.1 Introduction	134
5.2 Formative Review	135
5.3 Formative Review Stages	136
5.4 Designing the Formative Review Instruments	137
5.5 Participants' Sampling.....	137
5.6 Data Collection Arrangements	139
5.7 Data Collection Processes of the Formative Reviews	139
5.8 The Expert Review of Prototype 1: Sampling and Participants	139
5.8.1 The Instruments Used in Data Collection for Expert Reviews	140
5.8.2 The Experts' Feedback.....	141
5.8.3 The Changes Made to the Checklist Prototype 1 (Experts Reviews).....	143
5.9 One-To-One Review of the Checklist Prototype 1.....	145
5.9.1 Introduction	145
5.9.2 Sampling and Participants.....	145
5.9.3 Procedures and Instruments	145
5.9.4 Changes Made In the Checklist Prototype 1 (One-To-One Reviews)	146
5.10 Small Group Review of the Checklist Porotype 2.....	148
5.10.1 Procedures	148
5.10.2 The Small Group Review (Experts).....	149
5.10.2.1 Sampling and Participants.....	149
5.10.2.2 The Instruments Used and Their Purposes.....	149
5.10.2.3 The Experts' Feedback (Small Group Review)	150

5.10.2.4 The Changes Made to the Checklist Prototype 2 (The Experts' Sessions)	152
5.10.3 The Small Group Reviews (Teachers)	152
5.10.3 .1 Procedures, Sampling and Participants	153
5.10.3 .2 The instruments used in data collection	153
5.10.3 .3 The Teachers' Feedback From Small Group Reviews	154
5.11 Field Testing of the Checklist Prototype 3	157
5.11.1 Introduction	157
5.11.2 The Field Testing Sessions' Procedures and Participants	158
5.11. 3 Feedback Questionnaire (Field Testing)	159
5.11.4 The Observational Log (Time, Performance and Behaviour)	162
5.12 The Changes to the Teaching Materials Evaluation Checklist Prototype 4 (after field testing)	167
Chapter 6 Teaching materials Evaluation and Design-Based Research: Results and Reflections	169
6.1 Introduction	169
6.2 The General Processes of this Study within DBR Phases	171
6.3 Understanding the collected data and their analysis	176
6.4 Teaching Materials Evaluation Checklist Rigour	177
6.4.1 The Checklist Validity	181
6.4.2 The Checklist Practicality	182
6.4.3 The Checklist Clarity and Appeal	182
6.4.4 The Checklist Effectiveness	183
6.5 Triangulation within the Study	183
6.6 Developing Teaching Materials Evaluation Checklists: the Results	184
6.7 Developing Teaching Materials Evaluation Checklists: Reflections	192
6.8 Design-Based Research and Criteria of Success: Reflections	199
6.9 Design-Based Research in Educational Settings	202
6.10 Design-Based Research Limitations	205
6.11 The Study Limitations and Future Dissemination	207
6.12 Final thoughts	209
Appendices	211
Appendix A: Subject Matter Experts Short CVs	211
Appendix B: Authorization Letter	214
.....	214
Appendix C1: Experts Questionnaire for the conceptual framework	215

Appendix C2: Experts feedback on the conceptual framework	216
Appendix D1: Brain writing Sessions Sheet (Teachers)	220
Appendix D2: Brainwriting Sessions Sheet (Students)	221
Appendix D3: Instructions for the brain writing sessions	221
Appendix D4: Pictures from the brainwriting sessions	223
Appendix D5: Themes of the brainwriting sessions.....	224
Appendix E: Short survey of the institutional needs	226
Appendix F: Formative review stages and instruments.....	227
Appendix G1: Experts feedback Questionnaire on the conceptual framework	228
Appendix G2: Experts feedback on the Checklist prototype 1	229
Appendix H1: One -to-one review protocol	233
Appendix H2: One -to-one data recording log	236
Appendix H3: One -to-one Feedback (Teacher 1)	238
Appendix H4: One -to-one Feedback (Teacher 2)	241
Appendix H5: One -to-one Feedback (Teacher 3)	244
Appendix J1: Small group presentation Slides.....	247
Appendix J2: Small group review questionnaire (for teachers and experts)..	248
Appendix J3: Small group review feedback (teachers)	249
Appendix J4: Small group review feedback (Experts).....	252
Appendix J5: Small group observational log (teachers).....	255
Appendix J6: Small group observational log (experts)	255
Appendix K1: Instructions for field testing sessions.....	256
Appendix K2: Field testing questionnaire	258
Appendix K3: Participants Field testing feedback	259
Appendix K4: Field testing observational log notes	263
Appendix L1: Teaching materials evaluation checklist prototype 1	271
Appendix L2: Teaching materials evaluation checklist prototype 2	279
Appendix L3: Teaching materials evaluation checklist prototype 3	287
Appendix L4: Teaching materials evaluation checklist prototype 4	293
References	297

List of Tables

Table (1) Higher Education Institutions (Al Shmeli 2009 cited in Baporikar, 2012	19
Table (2) Foundation Programme structure	32
Table (3) Design- based research phases of the study	35
Table (4) Comparison between assessment and evaluation from Parker et. al. (2001)...	44
Table (5) Typology of curriculum representations (Van den Akker, 2007)	53
Table (6) Formative & Summative evaluations from Nation and Macalister (2010)	65
Table (7) Purposes of user-based evaluation and the recommended number of participants by Petrie & Bevan (2009).	138
Table (8) The study Formative review purposes, methods and instruments.....	174
Table (9) General Processes of the Study	175
Table (10) Criteria for high quality interventions (Nieveen 1999, in Plomp 2013: 29)	179
Table (11) Summary of Theoretical, Empirical and Practical Findings.....	191

List of Figures

Figure (1) Checklist Categories and their Relationships (Wingate, 2002).....	69
Figure (2) Reeves models	95
Figure (3) Generic model for conducting design research in education. © 2012 Susan E.McKenney and Thomas C. Reeves.....	97
Figure (4) Conceptual Framework of the Sources for the Textbooks Evaluation Checklist in the English Language Programmes.....	101
Figure (5) The Revised Conceptual Framework of the Sources for Textbooks Evaluation Checklist in the English Language Programmes.....	105
Figure (6) The needs analysis triangle based on West (1994)	119
Figure (7) General Sequence of Formative Review Techniques (from Tessmer 1993)	137

Acknowledgement

All the praises and thanks to Allah for guiding me to the right path whenever there were more than one decision to decide on, and whenever the obstacles seemed difficult to overcome all the way through this study. I am also endlessly grateful to my supervisor; Dr. Paul Roberts who was always supportive and who helped me on several levels through the long journey of my PhD project. I am indebted to him for believing in me despite all the difficulties; without his constructive supervision and constant support, I wouldn't be able to finish this thesis. I would like also to thank the thesis TAP members Dr. John Issit and before him Dr. Florentina Taylor for their insightful discussions during TAP meetings. Every meeting, every comment and every feedback from them helped to enrich my personal experience and improve my thesis content.

Many thanks to the experts who validated the conceptual framework and the first prototype of the evaluation checklist: Prof. Brian Tomlinson, Prof. Jayakaran Mukundan, Dr. Saleh Al Busaidi and Dr. Vahid Nimehchisalem. Many thanks also to the other experts who participated in small group reviews and field testing Dr. Awadalkareem Ahmed Alhassan from Dhofar University, Dr. Rais Attamimi and Dr. Mohammed Al raimi, from Salalah College of Technology.

I would like also to thank the Director General of Colleges of Applied Sciences in Oman Dr. Abdullah Al-Shibly, the previous dean in College of Applied Sciences in Salalah Dr. Zaid Zabanoot and the present dean Dr. Ahmed Balkhair, Mr. Shakir Hamid Al Hassan Head of English Department, Dr. Halima Al badwawai the Assistant Director General for Academic Affairs in Colleges of Applied Sciences and Dr. Mohammed Bakhit Kashoob & Mr. Tarek Fstnassi for helping in the checklist excel sheet design.

I am also grateful to the teachers and coordinators from the six colleges of Applied Sciences who participated in the study: Ahmed Ali Al-Qasmi, Qasim Salim Al-Washahi, Ameera Al-Nabhani, Abla Alkaabi, Mohammed Al-Omairi, Khadija alharthi, Manju Gujadhur, Dr. Huda Said Al-Habsi, Dr. Ibrahim Al Mounla, Dr. Susan Crosbie, Jose Chacko, Alcindo Medeiros, David Chubb, Ediphil Sibanda, Felicia Bittues, Naveena Yates, Ali Sharaf, Issam Abdelrahman, Mohammed BaQawir, Marwan Yafai, Pablo Candela, Richard Day, Robert Tovey, Russell Calvert, and my friend Khadija Al Amri for participating in the earlier sessions and for always having time for me whenever I

need to talk. Many thanks to the 24 students who participated in the brainwriting sessions from the College of Applied Sciences in Salalah.

I am also thankful to all members of my close and extended family who took care of my children while I was in UK.

Declaration

I declare that this thesis is a presentation of original work and I am the sole author. This work has not previously been presented for an award at this, or any other, University. All sources are acknowledged as References.

Muna Kashoob

Dedication:

For those who were always proud of me -who left our world-my mother Fatima 2014, my father Mahad 2012, my brother Bakhit 2013 and my nephew Khalid 2011 and for my supportive family, husband and my children: Ahmed, Leena, Maher, Marwah, Amr, Amal and my autistic son Mohammed living with hope that one day he will be able to read these words.

Chapter 1 Introduction to the Study

1.1. Study Rationale

This study was initiated to solve one of the persistent workplace problems in the Colleges of Applied Sciences in Oman. Looking at the two dilemmas, the selection and evaluation of teaching materials for the different proficiency levels in the Colleges' Foundation Programme, it was concluded that these problems needed to be investigated. As a matter of fact, "most problems in teaching a foreign language are linked to the coursebook selection process. Once a coursebook is chosen, few efforts are made to evaluate the effectiveness of the book" (Bülent 2006: 21). The characteristics of workplace problems are defined Richey & Klein (2007: 17) as being "recurring and common to many settings" and can be "viewed as basically solvable" and are "reflective of broad areas of current interest in the field". So, according to this definition, materials evaluation is certainly a problem that is recurring, solvable and comprising different areas of ELT contexts. This problem is explored next through understanding teaching materials difficulties in the English language programmes as well as their selection and evaluation.

1.1.1 Teaching materials in English language programmes

In any English language programme, the coursebook is a very important component whether it is printed or electronic and programmes in the Colleges of Applied Science in Oman, the case subject of this study, are no exception. Though some might argue that not all institutions depend on published materials, Gray (2010: 189) emphasizes that "the future of such materials seems secure". It is not only secure but also increasing, especially in developing countries. Garinger (2002: 1), postulates that "even with the development of new technologies that allow for higher quality teacher-generated material, demand for textbooks continues to grow, and the publishing industry responds with new series and textbooks each year". Indeed, teaching materials and textbooks are one of the main concerns of authorities, teachers and users in English programmes especially with the huge and increasing interest in English as a foreign and international language. For many institutions, "it is widely accepted that the coursebook lies at the heart of any English language teaching situation" as Tsiplakides (2011: 758) quoting (Sheldon, 1988; Hutchinson and Torres, 1994) viewpoints. Currently, the availability of English textbooks is not a problem as there are hundreds of textbooks available in the market. In fact, it can be said that "there is a wealth of EFL material available" McDonough, Shaw & Masuhara (2013: 51). Likewise, Harwood (2010: 205-206) states that the interest in "Global

textbooks (GTs)” is increasing and that they “have become a major feature of Teaching English to Speakers of Other Languages (TESOL) pedagogy in the 21st century”. He suggests that “it is the major driver of global economy” and gives an example that “at Cambridge University Press and Oxford University Press” and during “years of global economic austerity, internal reports reveal that annual sales in ELT textbooks and related learning materials have continued to increase by between 9 and 12 percent, and make up 40-50 per cent of their total profits.”. In this context, textbooks “may be loosely defined as a published book, most often produced for commercial gain, whose explicit aim is to assist foreign learners of English in improving their linguistic knowledge and/or communicative ability” (Sheldon 1987:1). Because of the massive role of the textbooks in the English language programmes, “an approved textbook may easily become the curriculum in the classroom" (Fullan 1991: 70, cited in Lamie 1999: 2). Not only that, but the selected “coursebook for an ELT programme” will become “the textbooks for the years to come” McDonough et.al (2013: 51) as it is the case in the Colleges of Applied Sciences in Oman. Hutchinson & Torres (1994: 315) state that “the textbook is an almost universal element of [English language] teaching” and that “millions of copies are sold every year” so “no teaching-learning situation, it seems, is complete until it has its relevant textbook.” In fact, “the reality of most ESL/EFL classroom settings provides clear evidence of a preference for teaching with textbooks” (Byrd & Schuemann 2014: 380).

Commercial teaching materials are marketed as the best possible options for educators as effective tools to teach and learn the English language. The potential users are expected to use them without any attempt to question effectiveness or practicality. Published materials do not provide any schemes for selection or evaluation that may enable stakeholders to think critically about teaching materials in English language programmes. Ignoring such a crucial aspect as teaching materials evaluation leads to the negligence of important issues such as the pursuit of course improvements and teachers’ professional development opportunities. Similarly, the false security that these textbooks offer for students, teachers and institutions, prevent them from looking at them analytically or seeking other alternatives as their main focus becomes how to obtain the latest versions and copies of newly-released materials.

1.1.2 Selection of coursebooks

Despite the abundance of such materials, users find difficulty in choosing the appropriate textbooks for their English language programmes in their institutions. They will have to choose from two options: to trial several textbooks series, which is very expensive, or to evaluate recommended titles using an evaluation tool. It is obvious that “wider choice means more need for evaluation” (McDonough et.al 2013: 51) .The most recent attempt in the Colleges of Applied Sciences to introduce new teaching materials was the trialing of the English File series. The procedures for selection and evaluation were based on recommendation from the Head of the English Department. This method of selection suggests that “materials are often evaluated in an ad hoc, impressionistic way, which tends to favour materials which have face validity (i.e. which conform to people’s expectations of what materials should look like) and which are visually appealing” (Tomlinson 2013 a: 5).The colleges purchased copies and tried them with three groups. Later, the opinions of the teachers and students about the effectiveness of the materials was sought. Eventually, the materials were found to be culturally inappropriate for the Omani context. Such problems can be avoided if there is a simple and practical instrument such as an evaluation checklist, which can save time and money. In institutions such as the Colleges of Applied Sciences in Oman, where “the context in which language is taught is crucial”, where most of the decisions are “operated in a bureaucratic and hierarchical fashion”, and where “individualism and creativity are not particularly important” (Bülent 2006: 27), developing an evaluation tool may lead to more participation in materials selection and evaluation by the involved stakeholders, particularly teachers and programme coordinators.

It is recognized that “many teachers have no voice in textbook selection if they work in settings where textbooks are selected through an administrative process” or “by the program director” (Byrd & Schuemann 2014: 384). In the English language programmes, “the textbook is a 'problem' evincing a complex of difficulties in its creation, distribution, exploitation and, ultimately, evaluation” (Sheldon 1987:1). Even though “textbooks are major sources of contact with the target language” and “selection is an important decision”, the “educators need to be systematic and objective in their approach, adopting a selection process that is open, transparent, accountable, participatory, informed and rigorous” (Meurant 2010: 89). As there have not been any criteria for textbooks selection or evaluation in these colleges, the current teaching materials in the Colleges of Applied

Sciences were designated by the programme director. They were thought to be the best appropriate coursebooks for the English Foundation Programme. The previous selections in these colleges were also based on impressionistic recommendations. The publisher then was contacted and the materials series were bought for the six colleges. These textbooks are not reusable, so they are bought for every academic year.

Some researchers call for more involvement of all stakeholders, including administrators, suggesting that “the textbook should be evaluated, not only by the teachers or critics but also by the educational administrators who are responsible for building up the best value of textbook” (Lee, 2013: 81). The designing of a practical tool or checklist means that it is understandable and easy to use by all users. This can only be achieved through considering various sources, as Mahmood (2011: 1) explains: the “quality of textbooks” is “based on social needs, overall educational objectives, and up-to-date pedagogical and psychological theories of learning”. Moreover, materials selection and evaluation can help teachers to “to analyse their own presuppositions”, “establish their priorities” and “see materials as an integral part of the whole teaching/learning situation” (Hutchinson 1987: 42-43). Such general foundations to evaluate or judge the quality of textbooks are mentioned in several sources, but they are never made known to the users or the designers of such evaluation instruments. As a result, many teaching materials evaluation instruments developers can create their own criteria with no obvious bases or specified frameworks.

1.1.3 Materials Evaluation and Language Programmes

In Colleges of Applied Sciences, there is no real evaluation of the effectiveness of materials and apart from quality audits that are too general and only focus on the programme alignment with the National Standards and general learning outcomes rather than teaching materials quality, there is no practical tool that can help the teachers and the administrators to choose and gauge the appropriate teaching materials for the students in the English Foundation Programme. Furthermore, the audits are based on the portfolio prepared by the colleges and essential information about some of the important aspects such as teaching materials and their evaluation is never dealt with in depth. As a result, the quality audits can never be considered the ultimate solution that can tackle and solve the dilemma of teaching materials evaluation and selection. In fact, the Higher Education Institutions in Oman are facing huge problems in choosing and evaluating the teaching materials for the Foundation Programmes as there is no real research about the whole

education setting of the students' levels, teachers' experiences or recommended materials evaluation criteria for these programmes. Hence, the need to create and develop guidelines or checklists for materials use, selection and evaluation in the Foundation Programmes becomes a must in this initial stage of accredited education in Oman. Nunan (1998 cited in Balachandran, 2014: 209) states that:

The selection process can be greatly facilitated by the use of systematic materials evaluation procedures which help ensure that materials are consistent with the needs and interests of the learners they are intended to serve, as well as being in harmony with institutional ideologies on the nature of language and learning.

As there are no specific frameworks for developing textbooks and materials evaluation instruments, many schemes and criteria are used by researchers and evaluators to assess and revise the teaching materials. McDonough et. al., (2013: 52) state that “there does not seem as yet an agreed set of criteria or procedures for evaluation”. In fact, there are no clear foundations or frameworks on the sources for evaluation checklists, which has resulted in confusion for their designers and users. The lack of clear design guidelines, especially for novices, gives the impression that evaluation models and checklists are developed for specific users who have enough background in second language learning theories and practice.

In the literature, studies on material evaluation can be based on various methods such as SWOT (Strengths, Weaknesses, Opportunities, Threats) analysis as in Wisniewska (2011), the ACTFL standards (The American Council on the Teaching of Foreign Languages) as in Alemi & Mesbah (2012), the CEFR inventories (Common European Framework of Reference for Languages), as in Karababa et.al. (2010), on critical pedagogy, as in Rashidi & Safari (2011) or on post method principles, as in Hooman (2014). As a result, the design and development of evaluation instruments and checklists have been based on the efforts of different researchers and teachers. Most of these evaluation instruments have been categorized according to general concepts about what materials should look like and what items are important for the end users. Some evaluation tools or checklists are very detailed and others are one page in length. Despite their varieties “none of these checklists has taken into account the cultural, social and even political particularity and peculiarity of the educational milieu in which teaching and learning occurs” (Shatery & Azargoon 2012:1).

Razmjoo (2012) used Kachru's (1992) and Kachru and Nelson's (1996) classification of international users of English, which they call “concentric circles”, as a basis for their teaching materials evaluation instrument. The circles are: the inner circle (Native), the outer circle (ESL) and the expanding circle (EFL) where the third circle “comprises countries in which English has various roles and is widely studied but for more specific objectives” such as the “need of English for “reading knowledge” and “for scientific and technical purposes” (Razmjoo 2012: 123). The result of this checklist is “six major categories” that include “language components; tasks, activities and exercises; language skills; teachers’ manuals; general considerations; and critical discourse analysis (CDA) features”. Apart from the last category “discourse analysis features”, the rest of the headings are the same as previous checklists. The researchers are accustomed to the same headings and items and seem never to think of any innovative ones.

Shah, et al (2014) consider “Bloom’s taxonomy (1956) of Learning Domains” the “most appropriate for the evaluation of the textbooks. They “evaluated the coursebook on different levels e.g. cognitive, affective and psycho-motor” (Shah, et al 2014: 104). Littlejohn's (1998) framework is based on “materials as a pedagogic device, that is, as an aid to teaching and learning a foreign language” which limits “the focus to aspects of the methodology of the materials, and their content” (Littlejohn 1998: 182). His framework also depends on literature and “draws extensively” on previous models that he mentions such as Mackey (1965), Corder (1973), Breen and Candlin (1987) and Richards & Rodgers (1986). All the proposed schemes are mostly concerned with theoretical aspects of materials evaluation as they do not consider the students’ and teacher’s needs. User usability tests and formative reviews are also overlooked in most materials evaluation studies and projects. The challenge which all the previous schemes avoided is to design a set of criteria that can be used by ordinary teachers, who have little or no background in educational research and its academic jargon as well as specifying the exact sources of their schemes in a clear and easy model or framework.

There are some researchers who were near to developing effective evaluation instruments and frameworks for general English materials. For example, Breen & Candlin's (1987: 13) instrument or guide was “divided into two phases” The first included “initial questions” such as “(a) what the aims and content of the materials are (b) what they require learners to do (c) what they require you, as a teacher, to do (d) what function they have as a classroom resource”. Phase two inquired about issues such as “subject matter

topics, themes, ideas”, which they call the “subsidiary questions” phase. AbdelWahab (2013: 59) also used the concept of phases where phase one focuses on the “review of related literature and studies that tackled the process of developing EFL textbook evaluative checklist”; phase two on data collection that included “critical feedback survey, semi-structured interviews and written comments” and finally phase three that included the practical testing of the evaluation checklist through a “single case study” of the checklist use by two users and the researcher himself. So AbdelWahab’s checklist is based on “refined checklists previously developed by different researchers” (AbdelWahab 2013: 57). Despite their similarities, the checklists developed are different as each one represents the ideas and the background of its designer.

Another elongated attempt was conducted by Mukundan (2009), who used a four phase procedure in developing his framework for his doctoral study. He based his whole checklist on Skierso's (1991) Evaluation Checklist. After his PhD study about designing a ‘composite’ material evaluation framework, Mukundan participated with others in developing a teaching materials evaluation checklist that is also based on the previous developed evaluation checklists (Mukundan et.al. 2011). This checklist was first developed in June by Mukundan et.al (June, 2011). In the same year, “a focus group study designed to further refine” (Mukundan et.al. September, 2011: 21) their checklist that is “previously designed” by the same researchers. The following year, two of the researchers (Mukundan & Nimehchisalem 2012) surveyed the checklist for further refinement and improvement through consulting 207 experts in teaching ‘English as a second language’. The following enhancement was done by Mukundan & Kalajahi in 2013, where “944 male and female English teachers” were asked to use the developed checklist to evaluate their English textbooks. Also, Mukundan & Nimehchisalem (2013) tried to involve 82 evaluators to know their views about the usefulness of their developed checklist. The results of that study, according to the researchers, were positive, as the “the respondents generally agreed that the checklist is a useful instrument” (Mukundan & Nimehchisalem 2013: 810). Mukundan, with Nimehchisalem in 2015, changed the developed checklist based on “the comments of a panel” of three experts “who were sent a copy of the old version of the checklist” for further refinement. The experts “commented on the comprehensiveness, importance and clarity of the domains and items of the checklist independently” (Mukundan & Nimehchisalem 2015: 761). Based on that, the researchers changed the checklist in order to make it more practical and comprehensive.

There is also a designed website called ELT-TEC, for the developed checklist, which they stated “is the first online checklist” for the purpose of “English language learning textbook evaluation.” Despite their great efforts, there is no mentioning of specific sources for their teaching materials evaluation checklist. It seems that the problem of specifying or suggesting the evaluation instruments’ sources and basis is never considered by those studies. Following the same structures and schemes suggests that most developers try to avoid creating different or innovative sources for teaching materials evaluation instruments. Moreover, it can be said that all the steps followed in Mukundan's thesis, the five subsequent studies, and the website, can be completed in one single enquiry or study through using an innovative methodology such as design-based research. Definitely, what is lacking in this extended experience (from 2009 to 2015) of teaching materials evaluation checklist design and development is the presence of clear sources or grounds as their checklist is based, like most checklists, on reviewing previous ones and then selecting one as a starting point. Users never know the design processes or the procedures followed to design the evaluation checklist.

More examples of such studies include Hussin, Nimehchisalem & Kalajahi's (2015) evaluation checklist, whose checklist is “developed in the light of a number of previous evaluation checklists” (Hussin et. al. 2015: 27). Their checklist is also developed through three phases, similar to Mukundan’s (2009) thesis & the teaching materials checklist developed and refined by Mukundan and others in (2011, 2012, 2013 and 2015). But each developer has based his/her checklist on what he/she thinks are the appropriate items in a pick-and-choose method from the available alternatives. Mukundan & Hour (2010), in their review of 48 checklists over four decades (from 1970 to 2008), state that “checklist developers never think alike” and that one category can be “emphasized in different checklists under different sections” (Mukundan & Hour 2010: 339).

All these researchers have exerted great efforts to create useful teaching materials evaluation instruments. Despite that, their teaching materials evaluations instruments lack clear illustrations of their development, use and refinement. Also, their criteria or checklists may not be applicable to some teachers in the English programmes especially those who do not have a degree in education or have not studied any ELT courses. Some English language programmes hire English native speakers even if their degrees are in other areas such as psychology, history or any other discipline. Some of the non-native teachers who graduate from colleges of education may also lack the required experience

to conduct materials evaluation using such instruments. McGrath (2013: 117) concluded his discussion on how teachers evaluate coursebooks with these realities:

- (1) Teachers do not always determine the textbooks they use;
- (2) selection processes tend not to be based on systematic examination of the materials;
- (3) for the most part, teachers would like to be more fully involved in selection decisions, but be provided with guidelines to support them in this.

To avoid such confusion in teaching materials evaluation, specific sources have to be identified and general guidelines of its development have to be drawn; these are the main aims of this study, besides the development of an evaluation instrument and identifying its reviewing and testing methods. In other words, all the previous schemes and models of teaching materials evaluation are missing four core things. First, specifying a clear framework or model of the evaluation tools sources; second specifying a robust method for the evaluation instrument design; third providing clear guidelines for the development and use of the evaluation instrument, and finally suggesting a clear method for the developed evaluation instrument validation and review.

1.2 The Need for a Feasible Evaluation

Considering the problems mentioned of teaching materials evaluation tools, it can be concluded that there are several reasons that necessitate the development of new teaching materials evaluation instruments. Ten reasons can be identified, that demand the development of a teaching materials evaluation checklist which can be used in the specified local context of the Colleges of Applied Science and other English language programmes. First, the paucity of evaluation studies in developing countries (as most projects and research are developed in Europe and America), is a major cause for the need to design a practical teaching materials evaluation tool for the English Language Foundation Programmes in Oman. Carden and Marvin (2012: 106-116) state that the methodologies used in these developing countries are divided into “adopted methodologies, adapted methodologies and indigenous methodologies” where they mostly can be considered “Context Sensitive Evaluation” done by institutions or centres rather than “individuals”. In English programmes, due to their sophisticated contexts and needs, “adopted” and “adapted” methodologies are the most prominent. It is much easier for researchers to depend on previous research and developed checklists rather than designing new ones. Going for ready-made options is not confined to evaluation, but it is also the norm in teaching materials selection. Second, many evaluation instruments that

are created by different researchers, based on various assumptions, rarely involve teachers in the design and review processes. These instruments, methods and models are included in teachers' programmes at colleges and universities. But as soon as the teachers have graduated and started their work, they seem to be engaged in a dull routine where they have to follow the same curriculum, the same assessment and the same methods of teaching. According to Byrd & Schuemann (2014: 385), one of teachers' "limitations in working with a textbook is that they do not see it as a whole and do not examine the textbook in detail before the first day of class". Additionally, "those teachers who rely most heavily on the textbooks are the ones least qualified to interpret its intentions or evaluate its content and method" (Williams 1983: 251) and they can only learn how to use and supplement when they are involved in teaching materials evaluation. Unfortunately, in real practice, teachers are not involved in the materials selection and evaluation, so they are disadvantaged from an essential process that will allow them to improve students' learning as well as their own professional development.

The third problem in most of the previous evaluation tools is the impractical nature of the available methods and checklists. In spite of the enormous number of produced tools and checklists in the literature, their usage and practicality is criticized by many researchers and users. For example, Mukundan & Ahour (2010) conclude their study of 48 checklists with recommendations demanding more "clarity, conciseness and flexibility" as many of them are "neither tested for validity nor reliability" (Mukundan & Ahour 2010: 348). Accordingly, "these results have not led to a wide use of the proposed schemes and checklists to carry out systematic and reliable evaluations" (Karamoozian & Riazi 2008: 18). Moreover, Ansari and Babaii (2002) have criticized many checklists saying that the evaluation of textbooks "has thus far been ad hoc, with teachers trying to make decisions based on such unreliable and simplistic criteria as appropriateness of grammar presentation" (Ansari and Babaii 2002: 5) or even based on the most reputable and widespread textbooks among educators. Developing teaching materials for English learners in different proficiency levels is not easy, and neither is their selection and evaluation. Besides, "having very little time to choose from a vast option which may be common, popular, cheap, or all at the same time and with the ongoing pressure from the representatives of the different publishers most of the textbooks are chosen with little or no evaluation" (Zahan & Begum 2013: 193). For such issues, the development of a quick and practical evaluation instrument for these English language programmes is crucial.

Fourth, the design and development of textbook evaluation tools have not been elucidated in a logical and well-defined way. Few models by Scriven (2000 revised in 2005 & 2007), Stufflebeam (2000), Tomlinson (2003) and Bichelmeyer (2003) are available, that give guidelines for checklists in general, but still the need for detailed materials evaluation and explanation of their practical processes for many users and evaluators of teaching materials is required. These endeavours describe the processes rather than a theoretical framework that explains the basis and sources of their design and development. Fifth, most evaluation methods and checklists, “are not piloted” (Karamoozian & Riazi 2008: 13), lack clear guidelines and have problems in their practical application. Sixth, a practical tool such as an evaluation checklist can save time, money and efforts for many stakeholders. As Garinger states: “the use of an evaluation procedure or checklist can lead to a more systematic and thorough examination of potential textbooks and to enhanced outcomes for learners, instructors, and administrators” (Garinger 2002: 2). Also, Kiely (1995) suggests that the importance of evaluation is providing “information for specific decisions” as well informing “coursebook choices”. For example, evaluations can help to provide “data on students’ preferred learning pattern” and informing “decisions relating to the setting up of IT and self-access resources” in addition to empowering “teachers to innovate in their classrooms, document these innovations and use them for professional development purposes” (Kiely 1995: 11). Another benefit of evaluation instruments is “ensuring that the needs and wants of learners are given careful consideration when choosing English language textbooks” through applying “a written checklist of appropriate selection criteria” (AbdelWahab 2013: 59). Moreover, there is “a limit to what teaching materials can be expected to do for us” (Allwright 1982: 9) and through the process of evaluation, stakeholders can identify the strengths to encourage and weaknesses to improve besides supplementing appropriately. Indeed, evaluation instruments can have multiple uses and purposes. They can be used as selection instruments, evaluation tools, or even an instructional tools for designing teaching materials or as a part of the course or the programme curriculum development and constant enhancements. Hutchinson (1987: 37-38) refers to the role and the importance of materials evaluation in English language programmes:

Materials evaluation plays such an important role in language teaching that its potential for influencing the way teachers operate is considerable. Materials evaluation can and should be a two-way process which enables

teachers not just to select a textbook, but also to develop their awareness of their own teaching/learning situation.

Seventh, the difficulty to develop in-house materials for many educational institutions, is another reason that requires the availability of a practical evaluation instrument that “establishes procedures which are thorough, rigorous, systematic, and principled” and that “ensure that materials are devised, revised, selected and adapted in reliable and valid ways” (Tomlinson 2003: 5) In addition to that, “the selection processes can be greatly facilitated by the use of systematic materials evaluation procedures which help ensure that materials are consistent with needs and interests of the learners they are intended to serve, as well as being in harmony with institutional ideologies on the nature of language and learning” (Nunan 1991: 209). The planning and writing of materials in the English Foundation Programmes in the Colleges of Applied Science involve plenty of time, knowledge and experience in different disciplines and current trends in research as such projects need “the expertise, time, and funding which only a consortium of universities could obtain” (Tomlinson 2012: 150). They require knowledge in Applied Linguistics in general, language acquisition theories, second language learning and teaching theories, curriculum studies, teaching methods, learners’ and teachers’ needs and their strategies of learning and teaching as well as evaluation studies and theories. Also, writing in-house materials is not only demanding and time consuming, but the final materials may also fail to cater for all students’ proficiency levels and needs. In a local attempt in the language centre at Sultan Qaboos University (the leading university in Oman), Al Busaidi and Tindle (2010: 148), in their investigation of in-house materials are concerned about the feedback and usefulness of the materials (that are based on the “discovery approach” in teaching grammar) for low level students:

The texts produced by “low” students raise doubts about the effectiveness of this approach for learners with very low level of language...whether a different approach to language would be more effective with these learners is hard to say at this point. Further research needs to be conducted focusing specifically on the needs and learning styles of the weakest learners.

The problem of the in-house materials developed at Sultan Qaboos University can be attributed to the way materials developers think of the young adults or adult learners who are “often 18 to 40 years old” according to Tomlinson (2008). It “was noticeable that

lower-levels books appear to treat the learners as being low level in experience, intellect and maturity” and in most cases “the topics tend to be trivial and the activities are unlikely to stimulate the learners to think or feel” (Tomlinson 2008: 30). Also, using the same criteria in developing materials for different proficiency levels in English language programs may not fit all learners so “strategies from different theoretical perspectives may be needed” according to Ertmer & Newby (2013). For instance “task requiring a low degree of processing (e.g., basic paired associations, discriminations, rote memorization) seem to be facilitated by strategies most frequently associated with behavioural outlook (e.g., stimulus-response, contiguity of feedback/ reinforcement)” Ertmer & Newby (2013: 61). Sometimes, it is also difficult for the authorities in English programmes to give the teachers the freedom to create their own materials as their teachers’ recruitment includes teachers with different backgrounds, experiences and teaching methods. Similarly, the need for standardization and consistency may force such programmes to choose commercial textbooks to make sure that all the programmes will have the same objectives, content and consequently the same assessment.

Eighth, with the difficulty of developing in-house materials, published materials are not always the ideal alternative for some institutions. The textbooks produced may have many problems and limitations when used in different contexts. Therefore evaluation “could prevent many of the mistakes which are made by writers, publishers, teachers, institutions and ministries and which can have negative effects on learners’ potential to benefit from their courses” (Tomlinson 2013a: 6). Harwood (2010) talks about textbook research on the three levels of “content, consumption and production” and he mentions “gaps in all areas”. For example, at the content level, “there is less analysis of local as opposed to global textbooks including the under-researched area of teachers’ guides” (Harwood 2010: 2-3). Some of the problems or research gaps associated with consumption of these textbooks by teachers and students include issues like the relationship between these materials and the teachers and how they use them over long periods of time and how students make use of these textbooks outside the classroom. On the production level, there are not enough studies that allow the consumers or users of these textbooks to understand their “design process” or “design procedures”. Again, in such an ambiguous situation, materials evaluation instruments can be used for several purposes. They can be used as tools to critically analyse and understand these published materials, to select the

appropriate ones for different contexts or to evaluate the while-used ones either for improvement or replacement.

Ninth, novice teachers may consider English textbooks the sole and perfect resource for teaching English. Richards explains that “inexperienced ESL teachers whose mother tongue is not English may tend to follow the textbook very closely, to be very uncritical of their textbooks, and to be relatively reluctant to discard sections of the book and replace them with other materials” (Richards 1993: 7). This distorted concept, or as Richards calls it, “reification of textbooks” may “result in teachers failing to look at textbooks critically and assuming that teaching decisions made in the textbook and teaching manual are superior and more valid than those they could make themselves.” Therefore, the use of evaluation instruments to assess the teaching materials will help teachers to discover their problems and consequently find alternatives and make use of other supplementary materials to enhance students’ learning and their own professional development. Also, “the ability to evaluate teaching materials effectively is a very important professional activity for all English as a Foreign Language (EFL) teachers” (McDonough et. al. 2013: 50). Other advantages are mentioned by Nunan (1991) as “the best commercial materials fulfil an important teacher education function, and remove much of the burden and time involved in creating materials from scratch.”

Finally, previously-developed evaluation tools and methods seem to be ignored by users due to their unattractive design and contents, which justifies developing a new appealing and practical tool. Ansari and Babaii’s evaluation checklist, which they believe is a step towards establishing universal criteria or schemes for textbook evaluation, is based only on “close scrutiny of a corpus of 10 EFL/ESL textbook reviews plus 10 EFL/ESL textbook evaluation checklists” (Ansary and Babaii 2002: 1). Karamoozian and Riazi (2008: 2) depended on “several available textbook evaluation checklists”. Ayatollah (2010) also used the existing textbooks evaluation criteria as well as the viewpoints of some experts. Rahimpour and Hashemi (2011) tried to write their own criteria, but then they simply tested them on school teachers. Williams’ (1983: 251) scheme is based on literature and “assumptions about teaching a second language to a set of linguistic, pedagogical, general and technical criteria” and though it is old compared to other checklists, it is more comprehensive and inclusive despite the lack of contextual and practical aspect in its design. Mukunan and others criteria mentioned above tried to create

a reliable checklist where they attempted to allocate more time to its design and revision. They explain their procedures in Mukundan & Nimehchisalem (2012: 1128) as follows:

The project commenced by a review of the available instruments (Mukundan & Ahour, 2010). In the light of the evaluative criteria in the available well-established checklists, the researchers developed a tentative checklist (Mukundan, Hajimohammadi, & Nimehchisalem, 2011). This was followed by a qualitative study in which a focus group, including six ELT experts, helped the researchers enhance the clarity and inclusiveness of the checklist (Mukundan, Nimehchisalem, & Hajimohammadi, 2011). Parallel with the focus group, a survey of a group of English as a Second Language (ESL) experts' views on the tentative checklist was conducted.

Regarding all of the above aspects of materials evaluation, it seems that there is no thorough investigation to produce unified frameworks and guidelines that can demarcate the teaching materials evaluation instruments, their sources, processes and validation. Few attempts through doctoral dissertations can be found trying to cover such complicated aspects about evaluation checklists, but with no consideration of using innovative methodologies, though some of these dissertations were conducted at universities that are considered centres for checklists development and evaluation studies as University of West Michigan. Though these studies are not old, but innovative methodology like design-based research is not adapted despite its suitability for investigating such topics and themes. For example, Schroeter (2008) used qualitative nonexperimental and exploratory method, Guidy-Oulai (2009) employed three phases data collection processes with no clear specification of the methodology, Walker-Egea (2014) utilised mixed method with four phases. Though the idea of applying more than one phase in these checklists developments, the methodologies selected do not facilitate or cover the complicated areas of designing these evaluation instruments. Most researchers will recommend the use of their own designed instruments, and others like Tomlinson would suggest for every evaluator to create his/her own criteria of evaluation as “there can be no one model framework for the evaluation of materials; the framework used must be determined by the reasons, objectives and circumstances of the evaluation” (Tomlinson 1999: 11 cited in Tomlinson, 2014: 26-27). Though Tomlinson's point of view may accommodate all evaluators, it will make the area of materials evaluation more ambiguous and muddled as there is no specific model to follow nor clear instruments to use by different practitioners.

1.3 Procedures of Developing Teaching Materials Evaluation Instruments

In order to involve the teachers and coordinators in the six Colleges of Applied Sciences with the development of the evaluation instrument, the coordinators were approached first. The researcher, having been a coordinator for the English Foundation Programme from 2005 till 2008, is aware of the people who are close to the real problems and difficulties in the Foundation Programmes. Thus, an informal discussion with six coordinators of the English Foundation Programme in Colleges of Applied Sciences, about two important issues (selection and evaluation of materials) helped to decide about the importance of such matter in these programmes. The coordinators mentioned some of their own criteria which they thought may affect the selection and evaluation of materials. These criteria included teachers' and coordinators' experiences, the detection of their students' needs on a daily basis through classroom interactions, the results of the students' exams, some criteria from the internet and mapping the programme objectives against the currently-used materials. They did not have a specific method for evaluation nor an official obligation to assess the materials they were using. When asked about the need for an evaluation tool or checklist, five of them agreed that the availability of a practical instrument is necessary for materials selection and evaluation in the English Foundation Programmes. Only one coordinator thought that an evaluation instrument might not be useful as the materials selection is usually done by the programme director. This coordinator thought of the evaluation instrument as a selection tool and did not think that it may have other uses. Some of the reasons they presented include the role of such tools in helping the teachers to evaluate teaching materials at the end of each academic year, drawing teachers attention to different teaching materials and their evaluation, helping teachers to have general criteria to use when needed and assisting the new teachers to acquire a proper method to judge different materials. Choosing to discuss the need for an evaluation instrument with Foundation Programmes coordinators is based on their important role in the Colleges of Applied Sciences as they are familiar with both the administrative and academic circumstances of the Foundation Programmes. Their roles require communicating the Ministry of Higher Education rules and instructions to the teachers and staff in the Colleges and vice versa. In other words, they are aware of the Ministry's policies as well as Colleges' and stakeholders' needs. So, as the Omani context is lacking practical teaching materials evaluation instruments, designing such tools is considered very important for these institutions. Most of the previous evaluation tools have been developed in different contexts from the Omani setting and these are described

by Gray (2010: 39) as “normative” tools that “reflect the beliefs of their writers about the nature and scope of language teaching and learning.”

As has been demonstrated in the previous section, the problem is not only in the genesis of evaluation methods, but also in developing inclusive tools or instruments that consider theoretical and contextual aspects in English language programmes. The inadequate evaluation instruments and processes in general English programmes is acknowledged by Ali (2010: 85) referring to the limited “studies carried out in English for general purposes” compared to English for academic purposes and other English language learning contexts. Most evaluation studies depended on three main sources: Applied Linguistics research, materials’ users (students, teachers & experts) or previous evaluation tools and checklists templates. However, as far as the literature is concerned, none of the former checklists or instruments tried to combine these sources together. By looking at the various evaluation tools, and considering materials development and evaluation, it becomes obvious that two aspects (research and setting needs) need to be tackled through a robust developmental research. Likewise, the development of the checklist along with the manual or the instructions of use have been confined to very few checklists, where these guidelines are not properly linked to the checklist development and use. This may discourage practitioners and teachers from using these evaluation tools.

There are other issues that should also be considered, besides purchasing commercial textbooks and evaluating them such as the required training for their users which is “crucial to curb the misuse of a coursebook”, which will lead to help teachers to “gain a better grasp of how to adapt the book to fit the needs of their particular classroom environment” (Bülent 2006: 27). In the literature, many sources and recommendations have been considered in developing checklists, but designing a solid conceptual framework for the main sources, detailed description of the processes, clear guidelines and a thorough checklist validation with potential users have never been tackled together in one study. In reality, the relationship between materials design, research and evaluation has not been clarified and the theories and principles of materials design and development are not used clearly. Previous checklists are rarely accompanied by clear guidelines for users, which may explain the lack of interest in them and also “most of the checklists are prepared in haste and their reliability is questionable” (Zohrabi 2011: 216). Indeed, what is missing in the available evaluation instruments is clear underpinnings, guidelines of use and usability testing and appropriate reviewing method.

1.4 Overview of the Context

Education in Oman has a very short history compared to other nations. Due to social and political issues, the majority of Omanis were deprived of any kind of formal education till 1970. Before then there were only three schools in the whole country with about 909 students and 30 teachers. In 1970, the present ruler, Qaboos bin Said, started his reign, making use of all the available resources to modernize and develop the country in all aspects of life including education. By 2008, the number of schools had increased to 1052 with 553236 students and 41988 teachers (National report of the Sultanate of Oman, 2008). Sultan Qaboos University was also the first public University in Oman, established in the capital Muscat in 1986. Within the last forty years, the educational system has gone through major changes and reforms in schools and in Higher Education institutions. English used to be taught in Omani schools from fourth grade upwards, but in 1998, the government decided to teach English from the first grade besides the other core subjects such as Arabic, Math, Science and Islamic culture. Regarding Higher Education, many new colleges and Higher Education Institutions (HEI) were opened in all regions of the country. The six Colleges of Applied Science are among 30 institutions under the jurisdiction of the Ministry of Higher Education, besides five Private Universities and 19 Private Colleges, two of which are designated as University Colleges as Table (1) illustrates (Al Shmeli 2009 cited in Baporikar 2012: 12).

Under the Jurisdiction of	Higher Education Institutions	No. of Institution
Independent	Sultan Qaboos University (Government)	01
	Colleges of Applied Sciences (Government)	06

Ministry of Higher Education	Private Universities	05
	Private Colleges, two of which are designated as University Colleges	19
	Total	30
Ministry of Manpower	Higher College of Technology (Government)	01
	Colleges of Technology (Government)	05
	Oman Tourism College (Private)	01
	Total	07
Ministry of Health	Nursing Institutes (Government)	11
	Health Science Institutes (Government)	05
	Total	16
Ministry of Defense	Academies / Training Centers (Government)	04
	The Command and Staff College (Government with restricted admission)	01
	Total	05
Ministry of Awqaf and Religious Affairs	The Institute of Shari'a Sciences (Government)	01
	Total	01
Royal Oman Police	The Royal Oman Police Academy (Government, with restricted admission)	01
	Total	01
Central Bank of Oman	The College of Banking & Financial Studies (Quasi-Government)	01
	Total	01
Total Number of Higher Education Institutions		62

Table (1) Higher Education Institutions (Al Shmeli 2009 cited in Baporikar, 2012)

In the early stages, Oman was following a curriculum system that was based on the Arab culture of memorizing and rote learning of subject content, where students' goal is only to pass their final exams. This type of teaching and learning led to the engendering of high school graduates who mostly lacked the important skills that help them in their educational and social life. Al-Harhi (2012: 114) explains this as follows:

Education in most the Arab countries, including Higher Education, is similar to what Freire (1970), calls the model of “banking education.” In this model, education becomes an act of depositing, in which the students are the depositories and the teacher is the depositor. Therefore, reforming educational systems to create new systems that are based on critical cultural literacy, balances between internationality and locality, and focus on all aspects of globalization is the only hope for the Arab countries, including Oman to deal effectively with globalization.

In Oman as in other parts of the world, “the social, cultural and, indeed, political dimensions of English language teaching and learning have been increasingly recognized in recent years” (Hall, 2011: 181). As a result, and apart from religious and Arabic majors, most of the Higher Education providers decided to use English as the medium of instruction in their programmes and specializations. The spread of language schools and programmes are becoming essential elements in modern education especially in Higher Education institutions. Eventually, by 2006 “about 90% of all students entering HE are involved in some form of Foundation Programs” (Oman Academic Accreditation Authority). This is due to the demands and needs of “both globalization and Omanisation” as those two aspects “reflect the local and international challenges facing the country” (Al’Abri,K. 2011: 499).

1.5 Education Policies, problems and Reforms

As a developing, Arab and Islamic country, Oman has many considerations in terms of its national identity and dealing with global changes and demands. This can be perceived in its educational policy and principles mentioned in the UNESCO International Bureau of Education (IBE, 2011: 2) which aim to “integrate the individuals’ intellectual, emotional, spiritual and moral development; to nurture the capabilities of individuals and groups and to develop the spirit of cooperation; to modernize the Omani society by teaching the required technical skills and the proper intellectual approaches... to achieve social and economic progress...to achieve national unity...and to revive the Arabic Islamic Heritage.”

In order to be able to benefit from modern technologies and innovations Oman is aiming to be part of the international world. The commercial aspect of getting the newest developments and inventions cannot be separated from its cultural and political ones. It can be said that globalization comes to developing countries in packages these days, not as separate pieces. Al’Abri (2011: 500-501) agrees with this when stating that the “new

policies of education in Oman regarding curriculum, testing, English teaching, life skills, work skills, and computing skills are without doubt reforms and policy responses to the pressures and discourses of globalization.” He concludes his paper with the following interesting comments:

- 1) International organizations are playing a greater role in shaping the context of national education policies and becoming influential policy agents around the world, particularly but not exclusively in developing countries.
- 2) International organizations ‘impose’ their policy interests, objectives and philosophies in developing nations via conditionalities linked with loans.
- 3) The discourses of knowledge economy, life-long learning, international testing and technology are found to be the main concern of education policy in both developing nations; these have thus become in effect globalized education policy discourses.
- 4) English as the medium of teaching is taken into consideration by almost all developing countries’ education policies; this situation is a response to the globalization of the economy, with English now as the dominant language of business and trade.

Keeping the balance between modernizing the Omani Society and the country’s national and cultural essence and spirit is not easy. Al-Harathi (2002) summarizes the situation, in the Arab world in general and Oman in particular, towards globalization where “both the full resistance and the full surrender to globalization are not possible options for the Arab countries”, suggesting “a middle course...in order to adapt the globalization requirements and at the same time attempt to keep the local culture” (Al-Harathi 2002: 112-113). One of the local labour market demands and globalization requirements in Higher Education, is the movement towards using English as a medium of instruction in Colleges and Universities, clearly exemplified in the Colleges of Applied Sciences in six different regions in Oman.

In this complicated context, with local and international pressures, several reforms have been introduced. One of these was the new Basic Education Programme, which was announced in 1995. The entire curriculum was based on a learner-centered approach with learners’ needs met and their “physical, affective, social and intellectual development” addressed (Ministry of Education 2001). By the end of grade 10, students should learn 4,500 words and by grade 12 students should know “6,000 to 7000 words to be adequately prepared for university studies” (Sergon 2011: 4). In fact, such decisions about types and numbers of words to learn in these language programmes has to consider many other

issues before making such decisions because the materials created and developed can either empower students or hinder their progress. At the university, students should sit for proficiency exams to determine their level of English. They will have two alternatives according to their exam results: to join their credited courses if their scores are high or to join the English Foundation Programme for one to two years to raise their English language proficiency levels and skills. In Sultan Qaboos University, the only public university in Oman, which accepts the best students with the highest scores in all the subjects and during “the Fall 2011 class—the third year of Basic Education graduates—some 2,451 students had to enroll in Foundation Year, leaving a mere three hundred or so first-years who were able to directly begin credit-bearing courses” (Daniels, cited in Sergon 2011: 5). Sergon’s study recommends many changes on different levels both within the Ministry and for the teachers. For example, some of his recommendations within the Ministry of Education are to: “hire more qualified teachers, have more realistic expectations (level 4 or 5 on IELTS is not sufficient for university level studies), involve teachers in the creation of the curriculum, and change the curriculum: make it more relevant, more palatable and more realistic within the time-frame of a semester” (Sergon, 2011: 30-31). For the teachers, the suggestions are to “make sure that they are doing their research and that they are always up to date on new theories of learning; always work to better themselves and to work harder to motivate students” (Sergon 2011: 31). All of these recommendations cannot be achieved unless teachers are involved in curriculum and materials development and evaluation on a regular basis. Making such major changes and overlooking a fundamental element in the curriculum and materials improvement, which is evaluation, will decrease the effects of these changes and reforms on students, teachers, institutions and the whole society.

Despite the reforms which require Omani students to study English language for twelve years on a daily basis in public and private schools, English is considered the main weakness for most students in their exit Diploma exams. Al Mahrooqi has done a study in which she investigated the low proficiency of Omani students who finished their secondary school education. She summarized their main problems as “ineffective teachers, inadequate curricula, uninterested students, limited exposure to English outside the classroom, unsupportive parents, a poor school system, and peer-group discouragement” (Al-Mahrooqi 2012: 263). In her study, the curriculum was the third cause of Omani school graduates' low proficiency in English, after teachers and students

themselves. In fact, these three elements (teachers, curriculum and students) form the main factors for the success of any educational programme. Unfortunately, the students' low English proficiency in school remains with them when they enter the college or university. As a result, most students struggle in both English Foundation Programmes and degree programmes as these colleges' courses require higher levels of English. In such a situation, more attention has to be given to the selection and the evaluation of teaching materials in the English language programmes. This can help all stakeholders in these programmes (students, teachers and authorities) to identify their problems and find the appropriate solutions for them. Thus, teaching materials evaluation can facilitate these changes and their sustainability and continuation.

1.6 Colleges of Applied Sciences in Oman

The influence of globalization in Oman includes reforms in general education as well as in Higher Education. Before 2005, the six Colleges of Applied Sciences (Sohar, Sur, Salalah, Rustaq, Ibri and Nizwa) were Education colleges for teachers. All the subjects were taught in Arabic and the students did not face any challenging problems in their studies. Generally speaking Graduates were not good in English, but they were good in their specializations. In 2005, the government decided to change the role of these colleges from education to applied sciences, in which new specializations that suit the Omani labour market and the international interests in science and technology were introduced. So it can be said that the move "from teaching English to teaching content" or "CLIL (Content and Language Integrated Learning)" (Hall 2011: 195) started in Oman through the Colleges of Applied Sciences. The Colleges' four specializations are: International Business Administration, Communication, Design and Information Technology. In these Colleges, the new majors require a new medium of instruction, which is English. Many English teachers have been recruited from different parts of the world through the Ministry of Higher Education and other private agencies. Computer labs have been installed and English books and a new curriculum have been implemented in the Colleges' English Foundation Programmes, as well as in the degree programmes. In order to prepare students for their majors which will be taught in English, greater attention is paid to the structures and the organization of these programmes in these colleges and other Higher Education Institutions. Despite their importance and the different plans to improve the function of the English Foundation programmes, most problems are still waiting for more practical and innovative solutions. Before exploring one of the solutions

(regular evaluation of teaching materials through a developed evaluation instrument), background information will be presented about the nature and the development of the English Foundation Programmes.

1.6.1 The English Language Programmes: an Overview

Harold Palmer was one of the early scholars who witnessed the advent of “the teaching of English as a foreign language to adult learners” (Howatt 1984: 213). Later, many great contributors have participated in researching and investigating English language curricula, teaching materials and teaching methods (Tyler 1942, Taba 1962, Stenhouse 1975, Rodgers 1982, Rodgers and Richards 1986, Dubin and Olshtain 1986, Yalden 1987, Brumfit 1984, Clark 1987, White 1988, McGrath 2002 & 2013, McDonough, Shaw and Masuhara 2013). Additionally, the foundation and huge expansions of organizations such as the British Council (1934), the Council of Europe (1949) and Teachers of English to Speakers of Other Languages (TESOL) founded in 1966 had participated in the research and spread of teaching English as a second and a foreign language. So the impacts evolved from the contributions of single researchers to university research centres and different countries collaborative projects. One of the great projects in the UK and Europe was the research on core inventories for different levels. This research resulted in the Core Inventory for General English as a part of the Common European Framework of Reference for Languages (CEFR) which “has been translated into 40 languages and is now accepted as the international standard for language teaching and learning” (British Council/EAQUALS, 2010). In America, the establishment of the Commission on English Language Program Accreditation (CEA) agency in 1999 was very important in assuring the quality of graduates of English programmes. In spite of the importance of such agencies, they do not provide detailed guidelines or standards for the curriculum and materials to be used in English programmes. Masuhara and Tomlinson (2008:35) differentiate between “general English in English speaking countries and English as a foreign language in non-English speaking countries” in terms of the materials developed for use in English language programmes. But despite the differences in the learning contexts and students’ needs and wants, the same coursebooks are used as the main source for learning and teaching in these different programmes. By evaluating seven students’ books used in GE and EFL, Masuhara and Tomlinson concluded that: “GE materials and EFL materials cater for different contexts”. Thus, “by trying to satisfy two different groups of learners, coursebooks seem unable to set clear objectives and to choose suitable

approaches. As a result, neither GE nor EFL users seem to feel that their materials completely satisfy their needs and wants” (Masuhara & Tomlinsonin, 2008: 35).

In Oman, educational institutions face the same difficulty in choosing suitable materials for students in the General English Foundation Programmes. These educational institutions are also involved in adapting the materials to their students' needs as well as to the accreditation organizations' requirements, which is what the HEPs in Oman are still struggling with at present. The Omani experience is more based on the American Agency (CEA), constructed on specifying general standards without any attempts to make use of the “Core Inventory for General English” developed by the British Council and the Council of Europe, or the experiences of local teachers in Oman. Also, the absence of needs analysis and lack of research projects and assessment surveys make their materials selection difficult and stressful as they become a matter of trial and error. The Ministry of Higher Education has endeavored, through the introduction of quality assurance and the issuing of national standards, to set broad guidelines to be followed in the Foundation Programmes in four areas of learning: English language, mathematics, computing and study skills. These general standards can help the Quality Authority to conduct its regular audits, but their actual implementation will need more research and investigation on the institutional and local levels.

1.6.2 The Foundation Programmes in the Omani Higher Education Institutions

With the increase of interest in English in the Colleges of Applied Sciences and other institutions in Oman, the English Foundation Programmes have become an essential element for students to prepare them for their higher degrees. General Foundation Programmes are the programmes that are introduced in the educational institutions that offer their degree programmes in English. They usually provide the students with the basics in English language, math and computer skills they need in order to pursue their Higher Education. In the Omani context, the General Foundation Programme (GFP) “is regarded as a non-credit program designed to academically prepare a student for their post-secondary studies” (Carroll, Razvi & Goodliffe 2009: 2). The general Foundation Programmes are basically based on intensive English teaching for two or more semesters or terms. Oman is also influenced by the international attention given to English as the dominant language of the latest developments in all fields of education such as science, math and technology. As there is no ready model for teaching English, which can be implemented in these Higher Education institutions, each one of them creates its

Foundation Programme which accomplishes its mission, vision and aims. Usually, these programmes are taught in the English departments or in separate units or centers that are committed to raising the students' competence in English to be able to succeed in their studies. These departments or centers are usually responsible for their own developments, innovations and changes internally. In recent years and with the appearance of quality assurance organizations where their motives and purposes include both accountability and the development of different products and programmes as explained by Weir & Roberts (1994), the evaluating and controlling of these programmes is left to the Quality Assurance Authorities. According to the Oman Academic Accreditation Authority (OAAA) specifications, a foundation programme is a course that has these characteristics:

- 1) It is a formal, structured program of study licensed in the Sultanate of Oman.
- 2) It is designed to prepare students for their postsecondary and Higher Education studies
- 3) It precedes the first formal year of Higher Education study.
- 4) It is only required for students who do not otherwise meet all the entrance criteria for the first year of their postsecondary and Higher Education.
- 5) It does not result in the awarding of formal academic credit to the student.
- 6) It is general in disciplinary scope, thereby preparing students for a wide variety of subsequent postsecondary and Higher Education program options.
- 7) It is not precisely 'Higher Education', but nonetheless falls within the ambit of the OAC.

(Oman Academic Standards for General Foundation Programmes n.d)

With such specifications, General Foundation Programmes play a great role in students' success and progression in their postsecondary education. They help the students to gain the skills required for their degree programmes, especially in English language. Therefore, National Standards have been issued to regulate these programme in all the Higher Education institutions. However, these standards are not enough to improve the learning outcomes of these colleges.

1.6.3 The Foundation Programmes National Standard

The Foundation Programmes are placed by OAAA in the Authority Manual under a broad title: "Student Learning by Coursework Programs". These programmes are audited and reported in a general way. There is no thorough investigation of the main items mentioned

in Oman Foundation Programmes Learning Outcomes Standards, which demands that students should be able to do the following by the end of the programme:

- 1) Actively participate in a discussion on a topic relevant to their studies by asking questions, agreeing/disagreeing, asking for clarification, sharing information, expressing and asking for opinions.
- 2) Paraphrase information (orally or in writing) from a written or spoken text or from graphically presented data.
- 3) Prepare and deliver a talk of at least 5 minutes. Use library resources in preparing the talk, speak clearly and confidently, make eye contact and use body language to support the delivery of ideas. Respond confidently to questions.
- 4) Write texts of a minimum of 250 words, showing control of layout, organization, punctuation, spelling, sentence structure, grammar and vocabulary.
- 5) Produce a written report of a minimum of 500 words showing evidence of research, notetaking, review and revision of work, paraphrasing, summarising, use of quotations and use of references.
- 6) Take notes and respond to questions about the topic, main ideas, details and opinions or arguments from an extended listening text (e.g. lecture, news broadcast).
- 7) Follow spoken instructions in order to carry out a task with a number of stages.
- 8) Listen to a conversation between two or more speakers and be able to answer questions in relation to context, relationship between speakers, register (e.g. formal or informal).
- 9) Read a one to two page text and identify the main idea(s) and extract specific information in a given period of time.
- 10) Read an extensive text broadly relevant to the student's area of study (minimum three pages) and respond to questions that require analytical skills, e.g. prediction, deduction, inference.

(Oman Academic Standards for General Foundation Programmes 2008)

These National Standards can form a national plan for the English Foundation Programmes. But, at the same time, the freedom given to Higher Institutions Providers to form their own objectives that are in line with the National Standards and the structures of the Foundation Programmes, as well as to choose their teaching materials and assessment methods, has resulted in having different experiences and consequently various views and practices in the General Foundation Programmes. For example, there are English Foundation Programmes that consist of 3 levels, 4 levels or even six proficiency levels. Teaching materials also differ from one institution to another. Some colleges use different textbooks, others use one course book along with in-house materials, and some use a package of multilevel series textbooks for different levels. As

a result, the work of the Oman Academic Accreditation Authority becomes more difficult and the audits may not reflect the real problems of the teaching materials in the English Foundation Programmes. Also, these audits every two years, the colleges need internal instruments to evaluate the teaching materials every semester and to select the appropriate materials every academic year. Consequently, teaching materials evaluation tools can make huge differences in the quality of learning and teaching in the English Foundation Programmes, beside their indirect effect on the professional development of the involved stakeholders.

1.6.4 Foundation Programmes Council Audits

The Oman Academic Accreditation Authority (OAAA) started conducting audits in the Higher Education institutions in the academic year 2009/2010. Five of the colleges which are the subject of this study have been audited. The reports of these colleges can exemplify and reflect the positive and negative aspects of implementing the national standards in the Foundation Programmes. Any study of the latest audit reports of these institutions, would show the neglected areas in the review or the inspection visits of the Academic Accreditation Authority. For example, one of the main points noticed in the Foundation Programme in Salalah College of Applied Sciences by the accreditation panel is that “over 60% of GFP students entered the lowest stream of the GFP and, according to interviews with staff, the Panel was informed that the drop-out from the GFP is between 30-40% on an annual basis” (Salalah Audit report, 2011). This means that the students did not cope with the general English Foundation Programme and the situation needs more investigation to understand their problems. One of the core causes of such failure is the lack of involvement of teachers and students in understanding, selecting, and evaluating the teaching materials. Providing evaluation surveys or instruments that enable learners and teachers to assess their teaching materials and specify their problems may help to solve their learning impediments. It is important to ensure “that careful selection is made, and that the materials selected closely reflect [the needs of the learners and] the aims, methods and values of the teaching programme” (Cunningsworth 1995: 7). Sur College recommendation emphasises “that the College of Applied Sciences Sur, in conjunction with the Ministry of Higher Education, review(s)” should make sure that “its Foundation Programme exit point so that it is in line with Oman’s Academic Standards for General Foundation Programme Standards” (Sur Audit report, 2013). Nazwa also received the same notification, which stresses that “consideration also needed to be given to the

appropriateness of the required exit level of students from its Foundation Programme to ensure that students are meeting the English language learning outcome standards” (Nizwa Audit report 2012). All these notifications will continue to appear with every accreditation visit unless these Foundation Programmes outcomes and standards are aligned with the teaching materials objectives and content, which can be assured through constant evaluation.

The importance given to auditing teaching materials in the Foundation Programmes is marginal compared to its importance in the whole programme. Also, the freedom given to the institutions in choosing the structure and materials for their Foundation Programmes does not mean that they can choose anything just for the sake of keeping the programmes running. On the contrary, they have to exert great efforts in choosing and monitoring their teaching materials. The expected learning outcomes are specified in the National Standards for Foundation Programmes Manual mentioned in section 1.6.3 above. They are "structured according to a revised formulation of Bloom’s taxonomy of educational objectives" (Bloom1956, cited in Carroll et.al. 2009). In spite of the efforts given to establishing the learning outcomes through international benchmarking and national consulting such as the “Report of the UK’s National Committee of Inquiry into Higher Education” and the Commission on English Language Program Accreditation (CEA), USA, the national two-day symposium in January 2007, and in post-symposium online discussion boards” (Carroll et.al.2009: 4), the experiences and actual settings of the institutions themselves were not involved. As Michael Scriven (in Alkin, 2013: 173) points out; “the quality control system is seriously flawed; it even lacks reliability, and there are also serious, unanswered concerns about its validity, for which at least follow-up studies are required, which are virtually never done”. As a result, the quality audits can never be the solution that can tackle and solve all the problems related to the teaching materials and their evaluation and assessment. And as Henson (2012: 5) concluded in a comparison of quality issues in both Oman and United Arab Emirates, “even today, several years on, institutions are still often unable to develop self-studies at a level that makes quality audit an appropriate mechanism”. The Foundation Programmes are considered a new experience in the Omani educational system and apart from a single reference that is edited by Saleh Al-Busaidi and Victoria Tuzlukova ‘*General Foundation Programs in Higher Education in the Sultanate of Oman: Experiences, Challenges and Prospects*’, there has not been any real attempt to conduct research in this area. The

teaching materials in these programmes are in themselves a substantial challenge for the colleges and universities, which most of them escape by opting for commercial course books. Fortunately in Oman, they have the budget to buy the published materials, but the problem remains unsolved: there is a lack of review, revision and evaluation of these materials in the Foundation Programmes on a regular basis which result in other problems in these institutions.

1.6.5 Foundation Programmes Structure in the Colleges of Applied Sciences

The Foundation Programme in the Colleges of Applied Sciences in Oman consists of four levels: A, B, C and D. The top level is A, and the lowest is D. There are four semesters and 20 to 24 teaching hours per week and the length of each semester is 15 weeks. Students sit for a placement test after their registration in the colleges in order to be placed into the right proficiency level. Students who are enrolled in D level (ENGL 3001), which is supposed to be equal to IELTS 2.5, study New Headway Plus Beginner. The students in Level C (4001) study New Headway Plus Elementary and they are expected to be at IELTS 3.0 at the beginning of the course. Both D and C levels courses focus only on General English Skills. In level B and A students study two courses: General English Skills and Academic English Skills. For level B (ENGL 5001 & 5002), which is equivalent to IELTS 3.5, the students are given New Headway Plus Pre-intermediate (level 1) for general English and Headway Academic Skills, reading, writing and study skills, as well as Headway Academic skills, listening, speaking and study skills for Academic English. The textbooks New Headway Plus Intermediate and Headway Academic Skills, reading, writing and study skills, along with Headway Academic skills, listening, speaking and study skills for Academic English are used with level A (ENGL 6001& ENGL 6002), which is equivalent to IELTS 4.0. The Foundation Programme, with its different levels, is planned to prepare the students for their degree courses (International Business Administration, Communication, Design and IT), which are all taught in English. The learners' English competence becomes essential as their subjects are taught in a language that is not used in schools.

In future, it is expected that Content and Language Integrated Learning, or CLIL “an approach to bilingual education in which both curriculum content– such as Science or Geography and English are taught together” (Graddol, 2008: 86) will replace other models of English language learning and teaching. CLIL along with teaching of English to Young learners and English as lingua Franca “will be the dominant trend and will

eventually replace English as a foreign language” (Morton, 2013: 111). Therefore all of the educational reforms and policies in Oman are attempting to conform to this concept of teaching English from the first grade to the establishment of the English Foundation Programmes in all the Higher Education institutions. Besides teaching Science and Math in English in some schools, the English Foundation Programmes can be considered the means to implement such changes and reforms in the Omani educational system. Table (2) illustrates the Foundation Programme structure in the Colleges of Applied Sciences.

Code	Name	Contact hours /week	Prerequisite
		ENGLISH & STUDY SKILLS	
ENGL 3001 (Level ‘D’)	General English Skills: Beginner	20-24*	None
ENGL 4001 (Level ‘C’)	General English Skills: Elementary	20	Completion – not passing – of ENGL 3001 or equivalent PT entry score
ENGL 5001 (Level ‘B’)	General English Skills: Preintermediate	10/11**	Pass in ENGL 4001 or equivalent PT entry score
ENGL 5002 (Level ‘B’)	Academic English Skills: Preintermediate	9/10**	Pass in ENGL 4001 or equivalent PT entry score
ENGL 6001 (Level ‘A’)	General English Skills: Intermediate	10/11**	Overall pass in ENGL 5001 + 5002 or equivalent PT entry score
ENGL 6002 (Level ‘A’)	Academic English Skills: Intermediate	9/10**	Overall pass in ENGL 5001 + 5002 or equivalent PT entry score
		COMPUTER SKILLS	
COMP 4001	Computer Skills: Basic	2	None
COMP 5001	Computer Skills: Advanced	2	Pass in COMP 4001 or equivalent PT entry score
		MATHS	
MATH 4001	Basic Maths	4	None

MATH 5001	Applied Maths	3	Pass in MATH 4001 or equivalent PT entry score
MATH 5002	Pure Maths	3	Pass in MATH 4001 or equivalent PT entry score

Table (2) Foundation Programme structure

Clearly, such fixed structure will need revisions annually, as “students' needs and interests are changeable by the time” and “a textbook works successfully today, it may not be so tomorrow” [sic] (Mahfoodh & Bhanegaonkar 2013: 3). Evaluation of teaching materials in the Foundation Programmes will enable the appropriate adaptation and supplementation that are needed for each level in every semester. As some “textbooks merely grow from and imitate other textbooks and do not admit the winds of change from research, methodological experimentation, or classroom feedback” (Sheldon 1988: 339), the need for continuous evaluation becomes a must in English Foundation Programmes. The power of evaluation may force publishers to consider research for different settings because they know that their published materials will be evaluated against specific criteria for selection and regularly while in use. In addition to that, “few teachers are really well-informed as how to choose textbooks and how to evaluate them properly” and teachers should “bear in mind that the aim of textbook evaluation is to find the best possible fit for a particular group of learners, together with potential for adapting or supplementing some of the material to make it more suitable and appropriate” (Kurniawan 2006: 3). This can only be achieved if there is a clear evaluation instrument that is utilised by these programmes and understood by the programme director, the teachers and the coordinators. This study aims to solve this problem through establishing and a robust method for the design of a viable and effective evaluation checklist that is based on a practical conceptual frameworks and an innovative methodology of design-based research. The phases of the checklist development will be elucidated in detail in the subsequent section.

1.7 The Main Phases of the Study

To avoid all the mentioned problems and obstacles in developing an evaluation checklist for the teaching materials for the English language programmes, design-based research methodology will be employed. This methodology is expected to connect the intricacies and complexities of teaching materials evaluation. All the aspects will be connected (evaluation instrument sources, its development, design guidelines, use and validation) through the use of this methodology within one study. Therefore, the developed teaching materials evaluation criteria will go through the three phases of design-based approach that is used in developing instructional tools and products. These phases comprise: the analysis & exploration phase, the design and construction phase and the evaluation and reflection phase as illustrated in Table (3).

Phases	iterations	instruments	Participants
Phase 1: Analysis and exploration phase	1.Rationale/ problem statement	Literature review and developing a conceptual framework.	The researcher
		Informal discussions	six coordinators in the six Colleges of Applied Sciences, Oman
		Expert appraisal of the conceptual framework for teaching materials evaluation checklist sources and basis (5 open- ended survey questions).	4 experts in teaching materials and evaluation checklists
	2.Colleges investigation of context and stakeholders' needs	Brainwriting	6 teachers and 24 students (from 3 proficiency levels)
	3. Detailed description of	Brainwriting data analysis	The researcher

	analysis and exploration phase	Participants feedback	6 teachers in Salalah College of Applied Sciences
Phase 2: Design and Construction Phase	4. Initial design of the checklist through 2 different prototypes based on the conceptual framework	The development of teaching materials evaluation checklist prototypes based on the results of analysis and exploration phase.	The researcher
	5. The merging of the two prototypes into one prototype	The two prototypes that are developed based on (the conceptual framework with its two main categories: research and needs analysis)	The researcher
	6. Creating the first design of the evaluation checklist (Evolutionary prototype)	The filtering of the combined prototype(s) to construct the main categories and sub-categories of the evaluation checklist	The researcher
	7. The developed evaluation checklist (complete prototype)	Developer screening through using 3 checklists on how to develop evaluative checklists (Tomlinson, 2013; Bichelmeyer, 2003 and Stufflebeam, 2000). Another checklist by Wilson (2013) is used later	The researcher
Phase 3: Evaluation and Reflection Phase	8. Expert Review	Expert appraisal through 4 open-ended questions	4 experts in teaching materials, their evaluation and checklists development

	9. One-to-one Review	Detailed protocol that included teacher' annotated comments, researcher's observation and debriefing questions at the end of the session	3 teachers to check clarity, appeal, errors, practicality and usability
	10. Small group review of the (complete prototype)	small group using a protocol that included short presentation, then using the checklist to evaluate the teaching materials and finally using feedback questionnaire at the end of the session	2 professionals in teaching materials and evaluation 6 teachers from Salalah College of Applied Sciences
	11. Field test (complete prototype)	Field test the finalized checklist in the six Colleges of Applied Sciences through face to face and online sessions. The instruments used are the checklist itself, an observational log and 6 open-ended questionnaire at the end of the sessions.	One expert from Dhofar University; 3 teachers and coordinators from Salalah Colleges of Applied Sciences and 6 Foundation programme coordinators and teachers from other Colleges of Applied Sciences

Table (3) Design- based research phases of the study

All the above phases will help in answering the four questions raised within this study. These questions inquire about the best method for teaching materials evaluation instruments development, trying to look at this topic from different angles to achieve practical and clear outcomes that can help stakeholders to understand its development and use besides recognising its critical role in the English language programmes.

1.8 Research Questions

Main question:

- How can an appropriate method be established to design and develop an evaluation checklist for teaching materials in English language programmes?

Sub-questions:

- What are the possible sources and basis for designing and developing teaching materials evaluation checklists?
- What are the design guidelines for the development of teaching materials evaluation checklists?
- How can teaching materials evaluation checklists be validated?

Chapter 2 English Language Teaching Materials Evaluation in the English Language Programmes

2.1 Introduction

In chapter 1, the problems of teaching materials selection and evaluation in the English language programmes were exemplified. Within that discussion, it was concluded that teaching materials evaluation processes are as complicated and important as their design and development. This is due to the various factors and disciplines that should be incorporated in their design, development and evaluation, in addition to the various needs of the setting and stakeholders. Indeed, materials evaluation can only be conducted appropriately when the teacher or the practitioner is equipped with both theoretical and practical knowledge. However, in English language programmes, these aspects are neglected as there are many issues to consider besides materials evaluations. To this complex situation, the available teaching materials evaluation models and types lack clear heuristics that can help guide developers, users and other stakeholders. Some of the models can be also difficult to use by teachers and other users in the English programmes. Despite the plethora of terms and the myriad of instruments, teaching materials evaluation is still unknown to some and vague to others. The main reason for such confusion arises from the detached methods and the undecided basis and sources of the evaluation instruments. In order to understand the theoretical and practical sources of the teaching materials evaluation instruments, this chapter will review materials definition and terminology as well as the essential role materials play in the English language programmes. This background about teaching materials will help in understanding how they are developed and consequently how they can be evaluated. Then, evaluation is explored with its various models, theories, types and methods and how it is used in teaching materials evaluation and revisions. The checklist method is discussed in detail as it is thought to be one of the most effective methods of teaching materials evaluation (McGrath, 2002). Finally, the sources of the teaching materials evaluation checklist design and development are demonstrated through a conceptual framework that can guide both the checklist development processes in this study and other developers as well.

2.2 An Overview of Evaluation Theories and Models

Evaluation as a field is still developing with its increasing importance these days. Its development went through seven main stages. Hogan (2007) gives a very inclusive and brief history of evaluation and the researchers who had great influence and who contributed towards its development. He thinks that “the historical development of evaluation is difficult, if not impossible, to describe due to its informal utilization by humans for thousands of years”. Hogan (2007: 3) explains the different stages of evaluation development using Madaus et al., (2000) who divided them into “seven development periods of program evaluation. First, the period prior to 1900, which the authors call Age of Reform; second, from 1900 until 1930, they call the Age of Efficiency; third, from 1930 to 1945, called the Tylerian Age; fourth, from 1946 to about 1957, called the Age of Innocence; fifth, from 1958 to 1972, the Age of Development; sixth, from 1973 to 1983, the Age of Professionalization; and seventh, from 1983 to 2000 the Age of Expansion and Integration”.

Though evaluation is used on a regular basis in many international institutions, it is still one of the most confused areas in terms of its methods and actual practice. Many researchers discussed that perturbed evaluation state. For example, Scriven (1996: 395) noted that “evaluation is a very young discipline - although it is a very old practice”. According to Conner, Altman, and Jackson (1984: 3), “evaluation is an established field and is now in its late adolescent years and is currently making the transition to adulthood.” Evaluation as a process can be considered one of the means used to judge and compare things in their different forms. The simplest example can be seen in the way people look at the advantages and disadvantages of items, decisions or even judging other people. It is also used in all disciplines as an instrument for their improvement. In Cambridge online dictionaries evaluation means “to judge or calculate the quality, importance, amount, or value of something”. Patton (1988: 301) defines the practice of evaluation as:

Systematic collection of information about the activities, characteristics, and outcomes of programs, personnel, and products for use by specific people to reduce uncertainties, improve effectiveness, and make decisions with regard to what those programs, personnel or products are doing and emphasizes (1) a systematic collection of information about (2) a broad range of topics (3) for use by specific people (4) for a variety of purposes.

The definition by Worthen and Sanders (1973: 19) is thought to be an inclusive explanation about the nature of evaluation in general. They indicated that any evaluation “includes obtaining information for use in judging the worth of a program, product, procedure, or object, or the potential utility of alternative approaches designed to attain specified objectives”. Worthen and Sanders (1973) evaluation definition is almost equivalent to the evaluation purposes in this study. This concept of evaluation is more inclusive and related to the processes of the development of the teaching materials evaluation checklist as well as the methodology used in this study. Its demarcation is compatible with the definition of teaching materials evaluation instruments developed in this study, which aims to achieve the same goals of the above definition. This definition also helps to focus the design of the evaluation checklist conducted in this study for both the review of the checklist developed and its potential use as an evaluative instrument that can be used “in judging the worth of a product” that is teaching materials while-use or their “potential utility” before use.

Alkin, Christina and Christie (2004 cited in Alkin, 2013:11-50), through their evaluation tree metaphor, try to explain evaluation foundations and purposes. They try to summarize the different issues and matters that accompanied the development of the concept over the years. In their tree metaphor, the roots of the tree or the basis of evaluation enquiry are initiated by three grounds: Social accountability, Social inquiry and Epistemology. First, the social accountability types are divided into goal accountability (government boards and upper levels of management); process accountability (whether reasonable and appropriate procedures for accomplishing those goals have been established) and finally outcome accountability which refers to “the extent to which established goals have been achieved”. Second, is the “social inquiry” root and its influences in methods and models of evaluation such as “experimental psychology”, quantitative sociology” and “anthropology”. The third root in the tree metaphor to clarify the evaluation field is “epistemology” with its various paradigms such as positivism, post-positivism, constructivism and pragmatism which influenced the “the use” of evaluation. The branches of the tree that emerged from the roots are also three, representing purposes, methods and values. The first branch is where the evaluators see evaluation as a tool to decision making (Stufflebeam as an example). The other two branches are methods which are “the techniques used in the conduct of evaluation studies” and valuing which is “placing the value on the subject of the evaluation” as “the evaluand, is essential to the

process” (Michael Scriven and Robert Stake are examples). Alkin, (2013) uses the word “theories” with caution when talking about evaluation as he thinks that “the terms approaches or models are more appropriate”. He then divides the models into two types: “Prescriptive models” as “a set of rules, prescriptions...and guiding frameworks” and descriptive models as “a set of statements and generalizations”. He stated that most studies are “predictive or offers an empirical theory” which means that they “have not been validated by empirical research”.

Evaluation models are tools to evaluate programmes and their different features. Scriven defines “models” as “a term loosely used to refer to a conception or approach or sometimes even a method (e.g., naturalistic, goal-free) of doing evaluation” Scriven (1976: 233). There are many models that can be used by evaluators nowadays, Stufflebeam (2001) names 22 models. He also presented an inclusive review of literature in evaluation models from 1942 to 2000 for the sake of conducting effective evaluations. Most of these models are basically used in companies and other organizations for improvement and accountability issues, and they require expertise, time and funding. Payne (1994) discussed the four broad divisions of the formative evaluation models which can encompass sub-models within them: these are management models, judicial models, anthropological models and finally consumer models. Brown (1995: 219-225) looks at the division of models in the “educational literature” to include also four broad groups: the “goal-attainment approaches, static-characteristic approaches, process oriented approaches and decision facilitation approaches”. In the first group “the focus of evaluation is on the goals and instructional objectives with the purpose of determining whether they have been achieved.” Examples of this group are manifested by the Tyler model in 1942, and the Metfessel & Michael model in 1967. The second group of models are “performed to determine the effectiveness of a particular programme” and the evaluation is “conducted by outside experts who inspect a program by examining various accounting and academic records, as well as such static characteristics as the number of library books, the number and types of degrees held by the faculty”. These models are usually used now in different institutions to inspect their effectiveness “for institutional accreditation” purposes. This type of evaluation has been conducted in different higher education institutions in Oman. But the problem of these evaluations is their general and broad focus, which outruns other important issues in the educational institutions. The third group is the “process –oriented approaches such as the Scriven model (1967) and

the Stakes model (1967), where “evaluators engage in descriptive as well as judgmental activities”. Finally, decision-facilitation approaches have a focus of evaluation which moves from goals and accreditation to making decisions. In this group, “the most important function of evaluation is to help in making decisions” and a famous example of this type of models is the CIPP model by Daniel Stufflebeam that was developed in 1971. The model’s acronym stands for “Context (rationale for objectives), Input (Utilization of resources for achieving objectives), Process (Periodic feedback to decision makers), and product (measurement and interpretation of attainment during and at the end of the program).”

These models are designed to be used in other disciplines, but can be used also in education, especially curriculum evaluation. They are mostly used in large evaluations by experts or by researchers. They are conducted over long periods of time and they can be also very difficult to use by practitioners. They are also not applicable for ordinary use by teachers and administrators for materials selection and evaluation as they need academic background, time, funding and training in order to be used in the evaluation appropriately. As a result, educators usually look for flexible and practical alternatives for these evaluations models. Some evaluation models developers realized that these models are not practical. For example, Stufflebeam turned his model content into a checklist called *CIPP Evaluation model checklist* to make it more applicable for the users. Checklists are considered the ideal substitutes for such models as they are the “distillations of lessons learned from practice” Stufflebeam (2002 cited in Wingate, 2002: 1). Accordingly, checklists become very important and effective tools in teaching materials analysis, selection and evaluation and can be applicable to most educators. Materials evaluation is not easy to conduct professionally as “it draws on two distinct and complex fields- curriculum and evaluation- both of which relate to dozens of different definitions, approaches and methods” (Levine 2002: 1). There have been many attempts to evaluate teaching materials in different programmes. Some researchers call them ‘evaluation tools’, others ‘evaluation types’ or even ‘evaluation approaches’. The confusion of terms is strongly evident in the evaluation of teaching materials as it is in the other types of educational research. Materials evaluation is considered a constituent of materials development as it is exemplified in curriculum design models (Tyler, 1942, Taba, 1962 & Brown, 1995). Certainly, the evaluation of teaching materials is not restricted to their development, but it is used in materials selection and the evaluation of

commercial textbooks to gauge their usefulness and effectiveness in different stages of their use. For evaluating the commercial textbooks, the only source for practitioners and evaluators is through the materials themselves. Hence, to understand the evaluation procedures, the instruments used and their purposes, the evaluators have to familiarize themselves with the various areas of their evaluation. In this study, the evaluation focus will be the teaching materials for general English in the Foundation Programmes which will be elucidated next.

2.3 ELT Materials Evaluation clarified

Alderson & Beretta look at educational evaluation and explain the reasons behind the “explosion of interest in programme evaluation. The first reason was the “launch of sputnik in 1957” which led to huge “federal funds in the US” in the development of “science curriculum”. The other reason is the “reforms of President Johnson in the USA” that resulted in the production of educational programmes such as “Sesame Street, Head start and Follow through” and all of these programmes needed to be evaluated “for the purposes of accountability” Alderson and Beretta (1992: 13). Likewise, the research of the famous educator Ralph Tyler that resulted in his book *Basic Principles of Curriculum and Instruction* in 1949 and his behavioural objectives had “major influence on evaluation” (ibid: 13). At the beginning evaluations were very systemic and objective which means that the evaluators checked the curriculum or the programme against the stated objectives. With the inception of constructivism theory in education, evaluation studies included qualitative approaches in their projects involving people, considering their opinions and their various contexts.

In the literature, there are many types of evaluation according to their aims and purposes, such as micro evaluations of a part of a programme and macro evaluations of the whole programme. Williams and Burden used “illuminative evaluation to gain insights into all aspects of the system in which the innovation takes place in an ELT project design”, (Williams & Burden, 1994). They differentiate between three types of evaluation: formative, summative and illuminative. The middle type between formative and summative evaluations, the “illuminative evaluation” is conducted when the “the evaluator is actually involved in the day-to-day working of the project, and thus attempts to shed light on issues that emerge”. Additionally, “the role of the evaluator in such an approach is to produce, not a set of ‘findings’, but instead ‘an interpretation of highly complex system” (Williams & Burden, 1994:23). In that sense, their illuminative

evaluation is considered a detailed description of the evaluation processes using different techniques. So, even with the occurrence of formative evaluations and summative evaluations, researchers like Williams and Burden thought that they are not enough research tools for their own study. Accordingly they thought of illuminative evaluation which had “arisen” from the work of Parlett (1981) and others who were unsatisfied with Tyler objective model and were searching for alternative detailed models for curriculum evaluation. The summative evaluation role is to evaluate the “final outcomes” and the formative evaluation “is ongoing in nature, and seeks to form, improve, and direct the innovation, rather than simply evaluate the outcomes”. Ellis (1997: 36) calls formative and summative evaluation “predictive evaluation and retrospective evaluation”, referring to their roles in an evaluation. Evaluation research has developed immensely, but in spite of that, it is still not fully developed. Currently, many terms are still confused with evaluation in different contexts such as analysis, assessment and audits. The confusion of evaluation can be clearly seen in the terminology used, types and mixing with other review tools. So, when conducting evaluations, the evaluators have to be careful not to confuse evaluation with audition or analysis. Tomlinson (2003: 16), for example, differentiates between ‘materials analysis’ and ‘materials evaluation’ as “evaluation focuses on the users of the materials and makes judgments about their effects” whereas “analysis focuses on the materials and it aims to provide an objective analysis of them”. Evaluation is also different from assessment in terms of its “purpose, depth of analysis and response” as explained by Parker et al. (2001) in table (4).

	Assessment	Evaluation
Purpose	To improve future performance	To judge the merit or worth of a performance against a pre-defined standard
Setting Criteria	Both the assessee and the assessor choose the criteria.	The evaluator determines the criteria
Control	The assessee --- who can choose to make use of assessment feedback	The evaluator --- who is able to make a judgment which impacts the evaluatee

Depth of Analysis	Thorough analysis by answering questions such as why and how to improve future performance	Calibration against a standard
Response	Positive outlook of implementing and action plan	Closure with failure or success

Table (4) Comparison between assessment and evaluation from Parker et. al. (2001)

Evaluation is also different from audit that is based on general standards (usually conducted by quality assurance agencies) where the institutions prepare their own portfolios to be inspected by the audit panels to make sure that the institutions are able to translate their mission and vision statements into practical applications as an assurance for accreditation. Accreditation here is defined by Brown (1995: 221) as “a process whereby an association of institutions sets up criteria and evaluation procedures for the purposes of deciding whether individual institutions should be certified (accredited) as members in good standing of that association.” Unlike audition, evaluation can be done by any of the stakeholders in the programme to decide about the value of the programme. Consequently, it is concerned with specifying the worth of the reviewed item rather than improving the performance, which is the purpose of assessments. Evaluation also differs from testing and measurement in terms of inclusiveness .As Brown (1995) explained, testing only refers to “procedures that are based on tests” whereas measurement includes testing and “other types of measurement” as attendance records, questionnaires, teachers rating of students” and similar tasks. Evaluation “includes all kinds of testing and measurements as well as other types of information, some of which may be more qualitative in nature” (Brown, 1995: 219).

2.4 Teaching Materials Definitions

There are various definitions of teaching materials in educational references. For Example, Harwood (2010: 3) stated that “materials is a term used... to encompass both texts and language- learning tasks: text presented to the learner in paper-based, audio, or visual form, and/or exercises and activities built around such texts”. Tomlinson (2013: 2) has his own definition to materials that “include anything which can be used to facilitate the learning of a language. They can be linguistic, visual, auditory or kinaesthetic, and they can be presented in print, through live performance or display, or on cassette, CD-ROM, DVD or the internet”. McGrath (2013: 2-3), introduces another broadened definition which defines teaching materials as:

- *A textbook, produced by a commercial publisher*
- *Commercial materials that are not provided as part of the textbook package...dictionaries, grammar books, readers and supplementary skills books.*
- *Teacher-prepared materials: selected by a teacher or a group of teachers...the authentic print or recordings like materials from songs, lectures, internet and You Tube...worksheets downloaded from internet, teacher-developed materials, games and realia (real objects used in the classroom).*
- *Any other visual or auditory means used in the classroom by teachers or learners (facial expressions, gestures, mime, demonstration, sounds).*

In all of the above explanations of teaching materials, the three definitions are concerned with the idea of being inclusive of all sorts of teaching materials used in delivering the lesson or the lecture in the classroom. The “gestures” and “facial expressions” are considered teaching materials. In such wide-ranging descriptions of a terminology, the basic knowledge and shared concepts lose their consensus among practitioners. As a result, teachers tend to use their own terms, methods and techniques, ignoring researchers’ findings and recommendations. Specifying a single and inclusive definition that can be shared by a majority of teachers will help to create a common ground between them and will encourage them to communicate, share and participate in research findings applications. It will also facilitate the teaching materials evaluation. As a practical definition of teaching materials, it is better to narrow them down to the actual materials that are used by the teacher effectively and successfully in the classroom. Some teachers will use only one type of materials and others will exploit as many types as they can. In both cases, the most important thing is not the range of the definition for the teaching materials, rather it is how the materials are used in the classroom. In this study the first two categories in McGrath’s (2013) definition above are the target for the developed evaluation instrument. Though the checklist developed within this study can be used for various teaching materials evaluation and selection, it is mainly for the evaluation of the first category -the materials that are designed for English language and learning contexts with their various components- the “commercial materials” as Nunan (1998: 208) call them.

2.5 Materials Design and Development

Materials development is defined by Tomlinson (2012: 143-144) to refer to “all the processes made use of by practitioners who produce and/or use materials for language

learning, including materials evaluation, their adaptation, design, production, exploitation and research”. Littlejohn (2012: 286) divides the development of English language teaching over three distinctive periods and traces the influences that affected each period by different language learning theories and consequently textbooks development and evaluation. For example, during the 50s and 60s, behaviourism “has been firmly cemented into language teaching materials, with its persistence right up to the present day, as the continuing use of drills, substitution tables and such like, demonstrates”. The next period was the 60s and 70s where the focus moved or shifted from the materials themselves to the students and how they can acquire the language, and resulted in “featured methodologies such as Gattegno’s Silent Way (Gattegno, 1972) and Lozanov’s Suggestopaedia (Lozanov, 1978)”. In the 1970s, for example, course designers “were influenced by the reconstructionist movement in general education and its arguments for objectives-driven courses” along with the “emerging perspectives on what communicative ability in language entails” (Hedge 2000: 346). So, the evaluators of teaching materials also changed their views according to the changes occurring in syllabus and materials development. Each period has its identified educational value systems (Clark, 1986) that varies in themes, topics and interests that affected foreign language curriculum design and evaluation. The 70s and 80s saw a huge leap, with the emerging of the communicative method and approaches where “functions and notions replaced grammar areas, and many published ELT course reappeared with a light dusting off to give it a new face, as grammar headings (such as The verb ‘to be’) were replaced with functional ones (such as Talking about existence)”, (Littlejohn 2012: 289). He argues that social and political notions as “McDonaldization and Neo-Liberalism” have affected the ELT field negatively. He criticizes the sequencing and fixed layout of the English language textbooks in which the order of items is almost the same where:

‘warm up’ activities may be routinely followed by some reading, which may be followed by grammar work, which may give way to written practice, before ending with some ‘freer work’ (the traditional PPP model, still being the norm). The precise nature of this sequence is not the concern here; rather, it is the fact that there is a fixed sequence, repeated across units, a proposed standardized ‘packaging’ of the classroom time. (Littlejohn, 2012: 291)

The development of teaching materials is based on different approaches and methods that emerged from language theories and models as well as theories of learning and teaching. For example “The combination of both theories of language and theories of learning led

to the production of teaching methods as in “the linking of structuralism (a linguistic theory) and behaviorism (a learning theory) produced audiolingualism” (Richards & Rodgers, 2014: 28). Such an interdisciplinary nature of research in materials development requires a broader framework for materials evaluation that incorporates both theoretical and practical aspects of language learning and teaching in English language programmes. Materials development and design for English language programmes is usually conducted by field experts and famous publishers as they require an academic background in second language learning and teaching theories besides practical experience. In doing so Finney (2002: 70) thinks that “the field of teaching English as a Second or Foreign Language (ESL or EFL) has largely ignored or been isolated from main stream developments, informed rather by research in linguistics and applied linguistics”. But, he then explains that this situation has changed over the years as “there has been an increasing awareness by ESL/EFL practitioners and theorists that indeed they are parallels (Stern, 1983; Richards, 1984; Nunan, 1988; Johnson, 1989) and that curriculum theory has much to offer.” Educational philosophies have always influenced the ontological and epistemological foundations of different curriculum theories and designs in different subjects including English language curricula. According to Finney (2002: 71), “three traditions are identified as Classical humanism, Reconstructionism and Progressivism, relating to the structural grammar/systems approach, the notional functional syllabus, and the process-procedural approach, respectively”. Also, Clark (1986: 94) summarises the influences of the three educational systems; the Classical Humanism, Reconstructionism and Progressivism in terms of evaluation, where it is determined by different evaluators according to each period. The shifts according to the above educational philosophies is clear, where the evaluation first incorporated “the inspectorate-both form outside the classroom” and “external evaluation to determine whether pre-specified goals have been achieved or not” to “the encouragement of teachers to evaluate their own classroom practices, and research their own solutions.” Richards & Renandya (2002: 73) also explain the influence of these philosophies on the development of ELT curricula where “the move away from grammatically based syllabuses in the 1960s led to a variety of syllabus proposals, including notional-functional, situational, lexical, task-based, and procedural, all of which claim to be examples of a communicative syllabus”. As a result, many models are created for materials development. These models, as introduced by Finney (2002: 71), are “the content model: Classical Humanism, the objective model: Reconstructionism, the process model: Progressivism” and finally “mixed –focus

curriculum: New Pragmatism.” So, syllabus types for the English language instruction are usually based on different teaching methods starting from the Grammar Translation method to the Communicative Method and finally to multiple methods that try to integrate any practical techniques that will help learners to acquire the language. According to Brown, (1995: 12), “mixed syllabuses occur when authors choose to mix two or more types of syllabuses together into what looks like a different type of syllabus-at least in the table of contents.”

Another movement that was initiated by Wilkins’ book the *Notional Syllabus* in 1976, with the focus on language functions, has influenced the new language inventories or lists that can be incorporated in English language materials development. For example, the Council of Europe “produced the syllabus specifications ‘Threshold level English’ (Van Ek, 1975) and ‘Waystage English’ (Van Ek, Alexander, and Fitzpartick 1977), which is widely used by textbooks writers” (Hedge 2001: 346) and which is based on “Wilkins’ notional syllabus.” Cunningsworth (1987) also traces the “influence of speech act theory (Austin, 1962; Searle 1969) on language teaching” which “made itself felt in the 1970s through the Threshold Level (van Ek 1975; van Ek and Alexander 1980) which set performance targets expressed in terms of the learner's ability to perform certain speech acts and to express certain concepts, or notions”. This movement led to many changes in our ideas about syllabus, curriculum and teaching materials development. Those inventories of language items and can-do statements were also criticised by some researchers. For example, Cunningsworth (1987:54) stated that “the speech acts included in the performance specification were loosely categorized and were called functions” and this resulted in the onset of “the notional/functional approach to language teaching”. These are developed now into the European Language Reference Framework which is becoming a branding hallmark for some published textbooks, like Headway and other published materials. Issues like the sequencing of the language items and the content of teaching materials units and activities have been debated by many researchers. So, instead of focusing, for example, on specific syllabus and items in materials evaluation, the evaluators encounter a myriad of complex issues that they should consider and include in the evaluation processes. As a result, “no single theory about language acquisition or language teaching dominates classroom practices”, so “we can expect language teaching materials to involve a variety of ideas about effective practices” (Byrd & Schuemann 2014: 381) in both teaching materials development and eventually in their evaluation.

One of the models for materials development is designed by Hutchinson & Waters (1987 cited in McGrath, 2002: 175) which consists of four elements for materials design: input, content, language and task. Although there is no agreement on materials development models, teachers and beginner developers can find a basis and guidelines for developing materials based on these available models (White 1988, McGrath 2002, Richards & 2013). Richards, for example, refers to two types of theories that materials writers are supposed to consult in order to create reliable materials. The first is “the theory of language and language use” in which the writer has to be familiar with the latest theories , for example, the “genre theory” that is “an interactionist view of second language learning” with “a systemic approach to grammar, an interactive model of reading, a task-based orientation to instruction”. The other theory is the “theory of language learning” which is very important as it “will determine how the syllabus is implemented in the form of exercises, tasks, activities and learning experiences” (Richards 2006: 3). In contrast to materials developers, the evaluators of the developed materials are faced with huge challenges as there are no specific models or frameworks that explain developing evaluation tools for teaching materials in English language programmes. Instead, they encounter copious methods, models and checklists that are difficult to apply in complicated educational contexts besides difficulties to identify their theoretical bases. To establish a model or framework for materials evaluation, three main aspects that are supposed to be manifested in the developed materials and which are important in materials development, and subsequently in their evaluation, will be considered. These aspects are the acquisition of linguistic items and how they are delivered through the teaching principles and pedagogical factors, and the way both linguistic content and delivery methods are put or organized together to form the ELT curriculum design. In order to understand the elements of the developed criteria for materials evaluation in the English language programmes, certain terms have to be clarified. The confusing questions for any published materials evaluators are what to evaluate and how. Should the commercial materials be called curriculum and evaluated according to its components? Or should they search for the types of syllabus underpinning the textbooks and consequently evaluate them against the syllabus principles? Or maybe the evaluators are just required to think about the physical items in their hands with their different names the teaching materials or the textbooks or the coursebooks (e.g. organization, layout, skills and sub-skills) as previous evaluation instruments suggested. As this study aims to produce and innovative tool, a short introduction about such terms will facilitate the

understanding and the design processes of the developed evaluation instrument. Before that, the role of the coursebooks in English language programmes is discussed.

2.6 The Role Of Course Books in English Language Programmes

Teachers and learners are concerned, in any educational setting, about the availability of English coursebooks. It is one of the most important elements, after teachers and students, for any English language programme. “In some places it is taken for granted that coursebooks are used as the basis for course” (Ur 2012: 197), the “hidden curriculum of the ESL course” (Richards, 2006: 1) or the “syllabus” (Harwood 2014: 1). In fact, the role of coursebooks in Colleges of Applied Sciences in Oman is a clear example of such dependence on published materials. Whether teachers are novices or experienced, the availability of “commercial materials” save their time and guide them throughout the whole course. In reality, many of the teachers in the English language programmes “have little or no formal teacher training” and thus “the textbook and the teacher’s manual are their primary teaching resources” (Richards 2006: 2). The debate about the role of textbooks is continuing, which suggests “a need to extend and strengthen the research base in this area” (Harwood 2005: 2).

To answer the question “Do learners need a coursebook?” Tomlinson (2001: 67) presents the arguments of the coursebook proponents as the following:

- 1) It is the most convenient form of presenting materials
- 2) It helps to achieve consistency and continuation
- 3) It gives learners a sense of system, cohesion and progress
- 4) It helps learners revise
- 5) It helps teachers prepare

The opponents of the coursebook say that:

- 1) It is superficial and reductionist in its coverage of language points and in its provision of language experience
- 2) It cannot cater for the diverse needs of all its users
- 3) It imposes uniformity of syllabus and approach
- 4) It removes initiative and power from teacher

An extreme position or opinion towards the advantages or disadvantages of textbooks is not the right decision. After all, perfection does not exist and everything including textbooks will continue to have its merits and demerits. Textbooks are helpful learning tools in the hands of students and teachers, but at the same time they could work as deskilling tools for teachers or restraining tools for students’ language proficiency as they

may prevent them from searching for new discoveries about language items and rules they are learning. Commercial English textbooks like any other subjects' materials may increase the "the problem of learning to learn" which lead to what Brown (1992) calls "diseases of schooling" that include "two stumbling blocks to lasting learning" the "inert knowledge" where "students acquire facts that they cannot access and use appropriately" and "passive learning" when "students do not really engage in intentional, self-directed action" (Brown, 1992: 144). Such problems can be overcome through good administrative rules which provide other learning sources for the students and constant professional development programmes for the teachers. The textbook then becomes an indispensable part that facilitates learning, but with appropriate guidance, supplementation and regular assessment and evaluation.

As Sheldon (1987: 10) state "the textbook is a 'problem' evincing a complex of difficulties in its creation, distribution, exploitation and, ultimately, evaluation." Regardless of criticism directed to commercial coursebooks, they are still extensively used worldwide. In the colleges of Applied Sciences, the Foundation Programmes are dependent on the multi-level coursebook New Headway Plus. In other institutions, in Oman, different coursebooks can be found based on the objectives and the plans of these different Higher Education institutions. Whether the programme curriculum is a coursebook, in-house materials or electronic resources in the future, its content will remain crucial for any educational context and though the form may change, the language content and items in the teaching materials will be the same in the current or future syllabuses. Therefore, developing a practical selection and evaluation instrument for this type of materials is the appropriate solution for both proponents and opponents. Because teaching materials evaluation instruments are the tools that will enable both parties to discover the advantages and make use of them, and the disadvantages, and supplement or replace them. Thus, the development of evaluation criteria in general and checklists in particular will continue to expand as they become the teachers' and institutions' devices to measure the effectiveness and the limitations of the teaching materials content as well as an effective instruments that can help in materials selection. Richards (2001 cited in McGrath, 2013: 9) assured that:

Teaching materials are a key component in most language programmes. Whether a teacher uses a textbook, institutionally-prepared materials, or his or her own materials, instructional materials generally serve as the basis for much of the language input learners receive and the language

practice that occurs in the classroom. In the case of inexperienced teachers, materials may also serve as a form of teacher training—they provide ideas on how to plan and teach lessons.

Even when some researchers reveal their dissatisfaction with the available textbooks and complain “about the published materials status quo” (Sheldon 1988: 3), the “coursebooks still continue to be the single most important resource in the language classroom throughout the world” (Arikan 2008: 2, cited in Tomlinson, 2013: 269). But again, the difficult questions to answer before attempting any teaching materials evaluation are how to conduct such evaluation and what elements should be evaluated.

2.7 How and what to evaluate?

Looking at textbook series or a coursebook, the evaluator may become bewildered with the amount of concepts associated with teaching materials, such as curriculum, syllabus, textbooks, coursebooks and teaching materials, which may look the same or slightly different. Misconceptions about some terms may confuse evaluators. For example, “in the USA, ‘curriculum’ tends to be synonymous with British understandings of ‘syllabus’ and, indeed, these two terms are often used interchangeably throughout ELT (including the UK)” (Hall 2011: 199). Therefore, and besides the above clarifications about teaching materials and curriculum, further moves which complement such design divisions would include providing a comprehensive instrument for teaching materials evaluation. Whether teachers call them course curriculum, syllabus, textbook, coursebook or materials, the purpose, which is to deliver knowledge and language input to the learners, will never change and neither will the evaluation goals.

Introductions to such terms can help in understanding curriculum and teaching materials and consequently make specifying the origins and components of the developed evaluation instrument clearer. There are various definitions of these terms in language teaching and learning and this may create confusion among many teachers and researchers. For example, “when there is a myriad of definitions of a concept in the literature (as with curriculum), it is often difficult to keep a clear focus on its essence” (Van den Akker 2007). Van den Akker agrees with Taba's definition, which suggests that the most obvious interpretation of the word curriculum is then to view it as a course, trajectory, or ‘plan for learning’ (Taba, 1962 cited in Van den Akker 2007: 37). He also differentiates between three main levels, as presented in the table below, “where the intended domain refers predominantly to the influence of curriculum policy makers and

curriculum developers (in various roles), the implemented curriculum relates especially to the world of schools and teachers, and the attained curriculum has to do with the students.”

INTENDED	Ideal	Vision (rationale or basic philosophy underlying a curriculum)
	Formal/Written	Intentions as specified in curriculum documents and/or materials
IMPLEMENTED	Perceived	Curriculum as interpreted by its users (especially teachers)
	Operational	Actual process of teaching and learning (also: curriculum-in-action)
ATTAINED	Experiential	Learning experiences as perceived by learners
	Learned	Resulting learning outcomes of learners

Table (5) Typology of curriculum representations (Van den Akker, 2007)

Nunan (1988: 3) quoting Candlin (1984) proposed that “a curriculum is concerned with making general statements about language learning, learning purpose, and experience, and the relationship between teachers and learners”. The word 'curriculum' can mean many things in different contexts and Su (2012) illustrates the different definitions in an algebra equations:

- 1) Curricula as a set of objectives = *goals or objectives*
- 2) Curricula as courses of study or content = *content + goals*
- 3) Curricula as plans = content + goals+ teaching methods
- 4) Curricula as documents = content + goals + methods + assessment
- 5) Curricula as experiences = content + goals + methods + assessment + extracurricular activities and learning environment + hidden curriculum + cultures (Su 2012: 153-154)

Brown (1995) also referred to these many concepts, describing the “sound teaching” activities that an educator or teachers have to be familiar with regarding the language curriculum. He acknowledged how Richards and Rodgers summarized the curriculum concepts into three broad and easy categories as “approach, design, and procedures” that were based on Anthony's (1965) concepts of “approach method and techniques”. This classification made thinking about ELT curriculum clearer for the educators as it was reduced into only three core components where the relationship between them is clearly illustrated and explained. He then introduced his own terms that included “approach, syllabus, techniques and exercises.” So thinking about curriculum within those specific approaches, can help educators develop their own concepts and knowledge about

curriculum and teaching materials and ultimately evaluate their effectiveness where 'approach' refers to theoretical underpinnings, the 'syllabus' to the general plan and 'techniques and exercises' to the actual application in the classroom. Based on the above-mentioned views about curriculum, including the ELT curriculum, it becomes clear that it is established on three main bases: theoretical or philosophical foundations, the actual implementation and the results or the outcomes of that enactment. As a result, any evaluation instrument designer should consider the ELT curriculum theory besides linguistic theories and pedagogical factors.

Richards (2001: 1) defines syllabus as “a specification of the content of a course of instruction and lists what will be taught and tested”, which resembles definition No.4 in Su’s equations of curriculum. So, it is a summary of the curriculum for the students, teachers and administrators. The final product or the curriculum can be offered in the form of several textbooks or a core coursebook. McDonough, Shaw & Masuhara (2013: 11) define syllabus as an “overall organizing principle for what is to be taught and learned...a general statement as to the pedagogical arrangement of content.”

Researchers also differentiate between textbooks, coursebooks and materials. Though each term has its own definition, when talking about materials evaluation in the English language programmes these terms will usually be used to refer to the same thing, which is materials used for learning and teaching. Sheldon (1987: 1) defines 'textbook' as “a published book, most often produced for commercial gain, whose explicit aim is to assist foreign learners of English in improving their linguistic knowledge and/or communicative ability.” So, a textbook, for example, is an educational book with the subject matter to deliver specific content, whereas a coursebook is defined by Tomlinson (1998: ix) as:

A textbook which provides the core materials for a course. It aims to provide as much as possible in one book and is designed so that it could serve as the only book which the learners necessarily use during a course. Such a book usually includes work on grammar, vocabulary, pronunciation, functions and the skills of reading, writing, listening and speaking.

Coursebooks and textbooks are almost the same except that a coursebook has the meaning of the core course materials. As mentioned in the section on materials’ definitions, Tomlinson (2003: 2) gives a very broad definition of materials to include:

Anything which can be used to facilitate the learning of language. They can be linguistic, visual, auditory or kinaesthetic, and they can be presented in print, through live performance or display, or on cassette, CD ROM, DVD, or the internet...

In curriculum and materials evaluation such background about these terms will help the evaluators and practitioners in their instruction and evaluations. And whether the evaluators will think of the teaching materials as a syllabus, a curriculum or a coursebook, a specific framework has to be identified that considers the core stakeholders and the various concepts that initiated the development of teaching materials. None of these important elements should be ignored when developing evaluation instruments. Therefore, to avoid the downsides of previous teaching materials evaluation tools and checklists, the sources that form the basis of the intended checklist are illustrated in a conceptual framework within this study. The purpose of this framework is to establish a clear structure about the disciplines, the categories and items that should be considered when it comes to teaching materials evaluation checklists design and development. It will help also any evaluation instrument developer to avoid the misconceptions associated with many terms used in the ELT field such as the ones mentioned above. Moreover, it will also form a clear starting point for the instrument development as well as the teaching materials evaluation.

Furthermore, in the mentioned situations and definitions above, it is clear that there are three main interested parties in curriculum development and evaluation: learners, teachers and policy makers. Thus, developing any teaching materials evaluation instruments should also consider them. Regardless of the concepts that are used as the basis for curriculum development, “materials are an important component within the curriculum, and are often the most tangible and ‘visible’ component of pedagogy” Nunan (1991: 227). Their evaluation “can be partly carried out outside the classroom” in form of “checklists and evaluative questions” and “in relation to real learners in real classrooms” (ibid.). Those terms may sometimes confuse teachers and materials’ users, but understanding them is essential in order to understand materials development and evaluation. For example, syllabus and curriculum models and approaches contribute to teaching materials development and consequently their evaluation. As said before, developing a viable and comprehensive evaluation checklist is supposed to be constructed on clear principles and guidelines that reveal the basis of materials design as well as their evaluation. In such situations, developing an evaluation framework that will identify the main sources of the

evaluation criteria, will enable the evaluators to focus on the process of evaluation instead of being lost in the morass of ELT terms and misperceptions. Though these terms have different definitions, most teachers consider them equal, which means that coursebooks, textbooks, syllabus, curriculum, and teaching materials may refer to the same thing for them. In this study, commercial materials, coursebooks and textbooks will be used interchangeably to refer to the English language teaching materials used in the English Foundation Programmes. So, an evaluation instrument that will be developed through this study, will mainly focus on three aspects from the literature: the theoretical foundations of the English language teaching materials, the teaching approaches and the ELT design, beside the contextual needs that will incorporate the needs of the learners, teachers and the authorities or the institutions.

2.8 Evaluation and Teaching Materials in English Language programmes

Evaluating English language materials in language programmes is a complicated process. The teaching materials evaluation instruments developers and users are supposed to be familiar with four main issues. The first, is the obscure and interdisciplinary areas related to materials development and evaluation which make identifying general guidelines for evaluation instruments development difficult and time consuming. The second is the problems of the evaluation instruments design processes and the differences that occur from the different viewpoints of different researchers and developers. The third is the actual use of such evaluation tools and establishing clear guidelines for use, and the fourth issue is the validation and reliability of such instruments. In this section, the difficulties of evaluation instruments design will be expounded, the approaches of several designers will be illustrated and examples of the actual use of such evaluation instruments will be demonstrated. The other aspects (design guidelines, guidelines of use and evaluation instrument validation) will be identified within the results of this study.

In the past, the methods used to evaluate teaching materials were easier and usually included “interview, personal visitation, questionnaire, and the study of catalogues and other published bulletins” (Harap 1934: 195). One of the early attempts on the evaluation for accountability, was by the American educator John Clement in 1942, who provided detailed descriptions on how different subjects in the American schools could be appraised or evaluated, including English. Later, many methods were discussed in the form of models, guidelines or checklists with and without rating scales (quantitative or qualitative). The previous methods and checklists were usually conducted through

choosing certain categories and items that the researchers thought were important for evaluating teaching materials. In most of these tools and checklists if not all, there are mostly missing elements that affect their practicality and usage. This can be the dependence on single source for development of the evaluation tools or checklists, the ambiguity of some items, or just the absence of a segment that deals with the trialling and the actual use of the teaching materials and the needs of their users. The evaluation content or categories and items are also varied according to different researchers. For example, the main categories that were suggested for English textbooks evaluations by Clement (1942: 67-76) were “the authorship of textbook and point of view, nature and organization of the content, suggestions on methods, instructional aids or helps for using the textbook, mechanical features including typography of the textbook and the format.” The evaluation instruments can vary in their categories and sub-categories, so designing a clear and practical guidelines or a conceptual framework for the evaluation instruments development will help to save time and to understand the design processes as well as the evaluation stages and procedures. Few endeavours by Tomlinson (2008, 2011 & 2013) have striven to establish checklists guidelines for their development and use in teaching materials evaluation. Roberts' (1996: 377) article *Demystifying Materials Evaluations*, was another attempt, where he introduces “materials evaluation as total process” that begins with pre-publication, through classroom trial and ends with decision stage. But this method needs time and expertise in order to be conducted.

Besides the differences in content and headings, the development of various English language syllabuses as mentioned above has affected the methods of materials evaluation from the “structural syllabus of the 1960s and the 1970s, the communicative syllabus of the 1980s, and the task-based syllabus of the 1990s” (Hedge 2000: 339) to the post-methods syllabus or the multiple or mixed syllabuses or “multi-dimensional syllabus” where the designer integrates several elements for the course design such as different situations, notions and topics, as well as language functions, structures and skills in the ELT materials development. Also, for the evaluation to “create ethos of openness, mutual respect, and trust” many institutions “prefer procedures which involve teachers” (Hedge 2001: 353) which means that the use and development of teaching materials evaluation instruments should consider the users’ needs besides the theoretical aspects.

McDonough, Shaw & Masuhara (2013: 24) discussed what is called a “multi-component syllabus”, indicating that “in the last 20 years or so... the idea of a multilayered-syllabus

has begun to be more explicitly and systematically addressed” which entails a multi-level tool for teaching materials evaluation as a result of these changes and expansions of the ELT curriculum. Researchers like Swan (1985), Harmer (1996) and McDonough and Shaw (1997), cited in Litz (2005: 24) “advocated an integrated, multi-skills syllabus because it considers and incorporates several categories of both meaning and form.” This new direction adds more complications to the evaluation instruments developers and users. In such a complicated situation depending on the same categories or terms or sources for materials evaluation becomes impractical for several reasons. First, evaluation instruments developers will keep changing the evaluation sources and terms. Second, the practitioners who are supposed to use these evaluation instruments will have to be familiar with all of these changes and backgrounds. Third, contradictions between the theoretical assumptions and their tangible representations (the teaching materials) may cause more confusion for the evaluators more than helping in judging the appropriateness or the effectiveness of teaching materials. As a result, this study aims to offer a new organisation of the evaluation sources, categories or checkpoints. This organization is not based on syllabus items nor curriculum categories such as content, objectives, skills, grammar, and vocabulary; rather it is based on general conventions about materials basis and sources. These are the two main strands: research and setting needs with their main headings and sub-headings that will be explained in details in chapter four.

In the literature, materials evaluations studies using evaluation instruments like checklists have been conducted by many researchers and the results of these studies indicates the importance of teaching materials evaluation instruments. For example, Litz (2005) used an evaluative checklist to evaluate a new textbook in South Korea while it was being used. The findings showed that the textbook “stood up reasonably well to a systematic in-depth analysis and that the positive attributes far out-weighted the negative characteristics” (Litz 2005: 34). In another study, Tomlinson et.al (2001) conducted an evaluation to review eight current adult courses issued by leading British publishers and they concluded that there are some positive aspects like the “qualities of many of the components of the courses” as well as the “attempts of many of the courses to encourage humour and fun” besides “the realism of many of the audio components of the courses”. Nevertheless, they recommended some reforms such as “a reduction in the attention given to explicit knowledge of grammar” and “an increase in the attention paid to helping learners...to achieve accuracy, fluency, and appropriacy” (Tomlinson et. al. 2001: 98).

They also encouraged the use of “literature and other genres which give adults something to think, talk, and write about” (ibid: 98). Tomlinson conducted a second study with Masuhara, Hann and Yi in 2008. The study revealed a “move towards stimulating more personal responses from the learners” “attempts to try to simulate real communication” and they were “impressed by the realism of many of the audio-visual components and by the use of the Internet”. However, they urged “materials producers to re-appreciate the value of the core student’s and teacher’s books”. They suggested “a teacher’s book which succinctly and clearly shows ways of effective and principled teaching that satisfies language learning theories” as this method can enable teachers to “personalize, localize, and adapt the global coursebooks to suit their learners in their classrooms...” (Masuhara et.al. 2008: 311). In 2013, Tomlinson & Masuhara conducted a third study where they stated that their “criterion-referenced prediction is that most of the courses” they “have reviewed, whilst being very appealing to the eye and to those users favouring discrete focus on and practice of language items, are unlikely to be very effective in facilitating language acquisition and development” (Tomlinson & Masuhara 2013: 252).

Another study by Ranalli (2002), on New Headway Upper-Intermediate, revealed that its “overall emphasis is clearly on oral communication”. This study could specify some of the advantages and disadvantages of this textbook. Some of the positives included comments indicating that “the language input is useful and relevant, especially the material focusing on the grammar of speech and vocabulary systems...” The negative comments or “the primary disadvantages lie in the methodology, which is somewhat restrictive and rests on some arguably shaky theoretical foundations.” This is because “the approach to accuracy work is rule-based and behaviorist”. Also, “it is questionable whether there is enough speaking practice of a useful type”. He concluded that generally speaking “the book’s faults are outweighed by its strengths and these can be overcome through adaptation and supplementation” (Ranalli 2002: 17). Johnson et al., (2008) study investigated “the textbook evaluation techniques of novice and experienced teachers, which was conducted in Lancaster University’s Department of Linguistics and English Language”. The implications of that study lie in its importance “in providing a research-informed basis for teacher training” (Johnson et al. 2008: 162), which means that such evaluation studies are much needed and are multi-purposed. They can inform the selection of teaching materials, the supplementation and the teachers’ professional development. As Wang (1998: 4) indicated, these criteria or “questions serve as guidance

for materials evaluators when they scrutinize a particular textbook they are using or they are going to select or adapt.” They also help teachers to “make decisions in selecting textbooks...form professional judgements” and “reflect on” their “teaching and learning experience.” With such advantages of conducting materials evaluation, greater attention has to be paid to developing appropriate instruments that can be used and understood by all the involved stakeholders in English language programmes. Rea, Dickins and Germaine (1994: 4 cited in AbdulWahab, 2013: 56) stated that “evaluation is an intrinsic part of teaching and learning” and Cunningsworth (1984 cited in Dougill, 1987: 29) assured that “professional judgement...lies at the base of the evaluation procedure.” These advantages and others were summarized as “some common objectives for evaluation” in general by Mahfoodh & Bhanegaonkar (2013: 4) to include:

- 1) determining the effectiveness of a particular intervention
- 2) finding out how well students are learning
- 3) identifying improvements which could be made to a specific course, learning activity or learning process
- 4) satisfying internal or external auditing requirements
- 5) demonstrating value to stakeholders (which might include project founders)
- 6) reflecting on professional practice in a structured way
- 7) building evidence for a portfolio (e.g. career development, teaching fellowship)
- 8) producing guidelines for colleagues (internal and external) who might want to carry out a similar innovation
- 9) generating data for a research study or publication
- 10) investigating an issue of personal, intellectual or professional interest

Attempting to connect all the interrelated issues of teaching materials evaluation, many practical evaluation tools, such as evaluation checklists, have been created for both selection and evaluation. The evaluation instruments development procedures involve considering the evaluators who will conduct the evaluation, the type of the instruments or the criteria that will be used, the users’ needs and the method for reporting the results of evaluation. These evaluation instruments can help to facilitate the whole evaluation processes and eventually yield effective and practical evaluations where the final reports can be used in the materials amendment, improvement and supplementation. Weir & Roberts (1994) differentiated between the “evaluation for *accountability* and evaluation for *development*” where the first examines “whether a course will be repeated...dropped or whether a particular source such as listening laboratory has been used sufficiently” and

the second “purpose of evaluation makes staff and/or institutions answerable to authorities and/or sponsor. It also makes publishers and textbooks writers accountable to teachers and teachers accountable to their students” (Weir & Roberts 1994: 4). In consequence, the instrument selected for evaluation is important to reach reliable results. Tomlinson (2013: 31), a key researcher in materials development and evaluation, recommends that “making an evaluation criterion–referenced can reduce (but not remove) subjectivity and can certainly help to make an evaluation more principled, rigorous, systematic and reliable”. He also suggests that “many of the checklists and lists of criteria ...provide a useful starting point for anybody conducting an evaluation”. Unfortunately, not all the instruments available for teaching materials evaluation are effective and practical in their use, as “some of them are impressionistic and biased” and “some of the lists lack coverage, systematicity and/or a principled base, and some give the impression that they could be used in any materials evaluation” (Tomlinson 1999: 11, cited in Tomlinson 2013: 35). The types and methods of teaching materials evaluation are discussed next to identify the most viable tool to be used in English language programmes.

2.9 Materials Evaluation and Types

There are many methods and types for materials evaluation. Some are more practical and easy to use and others need more expertise and longer time to conduct. As evaluation in general is clarified in section 2.2 through Brown (1995) and Worthen and Sanders (1973) definitions, teaching materials evaluation will be defined in this section. Tomlinson's (2003: 15) definition of materials evaluation is “a procedure that involves measuring the value (or potential value) of a set of learning materials. It involves making judgments about the effect of the materials on the people using them”. He advises the evaluators to: “apply” their own “principles of evaluation to the contextual circumstance” of their “evaluation in order to determine the most reliable and effective procedures”. Tomlinson suggests several principles for materials evaluation as well as the procedures and the items to be measured and tested. He proposes that through materials evaluation, the evaluator “tries to measure some or all of the following: the appeal of the materials to learners, the credibility of the materials, the validity, the reliability, the flexibility, the contribution made by the materials to teacher development the ability of the materials to interest the learners and the teachers, to motivate the learners..., the value of the materials in terms of short-term learning...and long-term learning...” (Tomlinson 2013 a: 21-22).

Despite his detailed clarifications, the evaluator may find difficulty conducting the evaluation for a number of reasons. First, the evaluation items are based on one source, which is second language acquisition principles. Second, there are no specific procedures for each item, and finally, there is no specific model or framework that can be used as a guide for inexperienced evaluators. Johnson (1989: 223) defines curriculum evaluation, which can be also used in materials evaluation, as “the systemic collection and analysis of all relevant information necessary to promote the improvement of the curriculum and assess its effectiveness and efficiency, as well as the participants’ attitudes within the context of the particular institutions involved.” The three definitions mentioned for general evaluation, curriculum evaluation and teaching materials evaluation share some general characteristics. These are judging the evaluand, identifying its effectiveness and utility to its users. Despite their comprehensiveness, it was thought that a more appropriate definition would include several additional issues in materials evaluation such as its goals, procedures, evaluators’ roles and reporting results. Thus, material evaluation in this study will refer to investigating teaching materials by any of the potential users for selection or improvement using a viable instrument leading to a final report about the evaluation procedures and results.

In spite of the huge number of studies that have been conducted on teaching materials, the checklist method instructions and items are still muddled, as is mentioned by many researchers including Tomlinson (2013) and Mukundan & Ahour (2010). Tomlinson (2012) presents a detailed literature review on how different researchers have proposed different schemes or criteria to evaluating teaching materials from Tucker (1975) to Ur (1996), Brown (1997), Hemsley (1997) and Gearing (1999). These criteria, contributed to the evolution of materials evaluation and understanding their content. However, there are still to be encountered many opinions and types of evaluations for teaching materials, which are mostly based on repeated procedures of designing and developing the pre-existing checklists and the evaluators' own experiences. Some of these methods are discussed in the subsequent sections as an introduction to specify the most practical and effective evaluation instrument for teaching materials in the English Foundation Programmes.

2.9.1 Macro and Micro Evaluation

Macro evaluations are inclusive compared to micro evaluations. In other words, macro evaluation is broader than micro evaluation. It involves administrative matters besides

the curricular matters. The curricular matters include materials, teachers and learners evaluations. According to McGrath (2002) the “macro dimension” has several stages which he calls the ‘approach’ and the “micro dimension occurs within each stage”. This micro dimension is the “set of technique employed”. Therefore, evaluators can do the two evaluations separately or together. Though this method seems comprehensive, it may not suit the contexts of the Omani English Foundation Programmes as it will require more time and expertise than what is available in the actual settings.

2.9.2 Pre-use, in-use and post-use Evaluation

McGrath (2002) puts the three types of evaluation in a cyclical approach to evaluate teaching materials. Cycle one is the pre-use, cycle two is the in-use and the third is the post-use. He states that the pre-use is the most popular type of evaluation as it is the easiest in terms of time and effort. The other two are difficult because “time is not available or has not been allocated” for them. Tomlinson has the same three types except that he substitutes the term “in-use” with “whilst-use” (Tomlinson 2003: 25-26). For him, the pre-use evaluation is “impressionistic” most of the time even if it is done against a number of criteria points. Most researchers use categories like the ones suggested by Tomlinson (texts, layout, tasks, objectives, local needs and pedagogical factors) to form the main checkpoints for materials evaluation. Teaching materials evaluation in-use, needs a close observation to all the details and activities that happen in the classroom. Though it takes more time, its results help considerably in materials development, improvement and in providing the appropriate supplementary texts and tasks. The post-use evaluation of materials according to Tomlinson can answer questions about different stakeholders in the educational institution. First, the students’ evaluation questions should enquire about the learners’ skills or abilities “which they did not know before”, the things that the learners can do “which they couldn’t do before”, how the materials prepare them “for their post-course use of the target language” and the effect they “had on the confidence/ the motivation of the learners”. Second, teachers’ questions in the evaluation examine if they find the “materials easy...and helped them to cover the syllabus”. Finally, the context is also covered through administrators’ needs and if they find materials a good apparatus that helped “them to standardize the teaching in their institutions” (Tomlinson 2017: 26). This method is clear but not easy to explain and conduct in the English Foundation Programmes in Colleges of Applied Sciences. So a new conception of evaluation that incorporates both views is essential. Considering the scattered views, all

the stages referred to by McGrath and Tomlinson above can be easily assembled into an easy and practical evaluation instrument.

2.9.3 External and Internal Evaluation

External evaluation is the quick and general look at the materials or the coursebook. In their evaluation model that includes external and internal reviews, McDonough, Shaw and Masuhara (2013: 59-60) focus on two things regarding external evaluation: the “blurb”, or the claims made on the cover...and “the introduction and table of contents”. After this stage, comes the second one which is the internal evaluation. The internal evaluation is a thorough examination of the course book content. The investigation includes the: “the presentation of the skills in the materials and the grading and sequencing of the materials”. The final step in this set of criteria is to check factors as “usability, generalizability, adaptability and flexibility” (ibid: 60-61) in the teaching materials. This framework seems similar to Tomlinson and McGrath’s pre-use and whilst-use evaluation, which includes both impressionistic and close evaluations. This framework is also comprehensive but not practical for the busy practitioners in the English Foundation Programmes.

2.9.4 Formative and Summative Evaluation

Since their coinage by Michael Scriven in 1967, these two types of evaluation have been used in evaluating different aspects of programmes and organizations. The two types can be used also to evaluate course materials and curriculum. Both methods with their techniques are used in conducting research projects. According to Stufflebeam & Shinkfield (2012: 314), Scriven’s formative-summative approach designates evaluation as “the systemic and objective determination of the worth or the merit of an object”. The “object” here can be anything including teaching materials. So in relation to materials evaluation, it can be said that formative evaluation is somehow equivalent to the in-use evaluation and that summative evaluation corresponds to the post-use evaluation. Nation & Macalister (2010: 126) compare the two types in terms of purpose, data type, data use and the “presentation of findings” in evaluating any course, as in table (6). Despite the importance of such types of evaluations, their use in the English language programmes and in other similar evaluations is not effective, hence, we can find researchers such as Williams and Burden (1994) looking for another complementary evaluation type such as illuminative evaluation, discussed earlier in section 2.3., which enables the researcher to

collect more data during the implementation of the intervention and find solutions for the problems that occur during the project enactment.

	<i>Formative</i>	<i>Summative</i>
Purpose	Improve the course	Judge the course
Type of data	More likely to look at causes, processes, individuals	More likely to look at results, standards, groups
Use of data	Used for counselling, mentoring, professional development, setting goals, adapting material	Used to make decisions on adequacy
Presentation of findings	Presented to and discussed with individuals	Presented in a report

Table (6) Formative & Summative evaluations from Nation and Macalister (2010)

As can be seen, different researchers and evaluators have their own terms, methods and models for materials evaluation. Despite the comprehensiveness of some frameworks, they all lack a focus on practical application in a complicated educational setting such as the English Foundation Programmes in Colleges of Applied Sciences in Oman. Varieties of terms and models sometimes complicate the situations rather than simplifying them, especially for busy practitioners who do not have time to read or understand all the issues related to the use of such schemes. A checklist that is divided into two parts (representing quick and detailed evaluations) which can save the time and effort needed to conduct such evaluations is a recommended alternative. In fact, many of the complications and unnecessary intricacies can be avoided through incorporating both types (quick and detailed evaluation) in one practical evaluation instrument such as an evaluation checklist designed specifically to be used in the English language programmes. So, the search for more practical and simple methods continues in the next section.

2.10 Practical Methods for Teaching Materials Evaluation

Beside the different types that can be used as evaluation tools for teaching materials, as well as other aspects of the educational programmes, there are particular types that can be used specifically to evaluate the different textbooks that are selected or will be purchased for the English programmes. To avoid confusing methods and models discussed in the previous sections with the suggested ones, the three methods are illustrated in the next sections. The three methods are the impressionistic, the in-depth

and the checklist. These methods are simple and easy to use by most stakeholders and they require less time and effort.

2.10.1 The impressionistic method

The impressionistic method is similar to McDonough et.al (2013), external evaluation mentioned earlier. It is a quick glance at the exterior features of the coursebook or the materials. All the descriptions on the cover of the book, the shape the size, the colours and the layout can be noticed and seen through the impressionistic method. McGrath (2002:25) expands the process to include “skimming through the book looking at organization, topics, layout and visuals”. But, as the name of this method suggests, it is not considered a comprehensive tool to judge the effectiveness of the materials. So it should be used along with another instrument or as a part of a comprehensive one.

2.10.2 The in-depth method

Both the impressionistic method and the in-depth method look at the claims made by the author. But the “in-depth method” is more detailed and thorough than the impressionistic method. It involves using many techniques that are mentioned by different researchers such as: a focus on specific features (Cunningsworth 1995), a close analysis of one or more extracts (Hutchinson 1987), or a “thorough examination of several units using predetermined questions” (Johnson 1986, cited in McGrath, 2002: 28). This method has disadvantages that are mentioned by (McGrath, 2002: 28) to include the following:

- 1) Representativeness of samples: the samples...selected for analysis may not be representative of the book as a whole...
- 2) Partiality: It gives only a partial insight into what the material offers.
- 3) Time and expertise required: some proposals for in-depth evaluation would involve a good deal of time; others require expert knowledge.

2.10.3 The checklist method

Checklists can be used in all disciplines as they are a very practical tool for all stakeholders. Scriven (2007: 1) defines checklists as “a list of factors, properties, aspects, components, criteria, tasks, or dimensions, the presence, referent, or amount of which are to be considered separately, in order to perform a certain task”. McGrath (2013: 182) stated that “even when teachers have the freedom to make the choice, either collectively or individually, they seem not to approach this in a systematic way” so he proposes that “one of the strongest recommendations in the professional literature is that a checklist be used.” He also stated that the reasons which discourage teachers from doing materials

evaluation are “lack of time, lack of training, and lack of confidence.” Therefore, in order to ensure a reliable evaluation, the teacher-evaluators have to be trained and provided with an appropriate instrument even if they are experienced. McGrath (2013: 124) advises that “while experience ...is valuable, it is not a substitute for training in evaluation.” Wilson (2013: 13) clarifies checklists general strengths as “they are easy to administer and use, less training is required than with other methods, the output produced by checklist is immediately useful, they can serve as memory aids” and finally their flexible nature that makes them easy to “customize” through “adding or removing sections or modifying items” makes their use easier than other evaluation instruments. McGrath also (2002) refers to the advantages of checklists in evaluating teaching materials as they are “systematic...cost effective... explicit...permitting a good deal of information to be recorded in a relatively short space of time” and “information is recorded in a convenient format, allowing for easy comparison between competing sets of material”. These checklists also offer “a common framework for decision-making” (McGrath 2002: 26-27). Scriven (2007: 4) also states the importance of checklists as tools that are having the characteristics of being “mnemonic devices” and “easy instruments.” They also help to “reduce the influence of the halo effect, i.e., the tendency to allow the presence of some highly valued feature to over influence one’s judgment of merit” besides reducing “the influence of the Rorschach effect, i.e., the tendency to see what one wants to see in a mass of data...”. Lastly checklists “can contribute substantially to (i) the improvement of validity, reliability, and credibility of an evaluation; and (ii) our useful knowledge about a domain.” Despite their importance, most developed checklists never explain how they are developed apart from the authors' descriptions of reasons for choosing the items to include in their checklists. Looking at the literature of checklists and their development, some instructions can be found for general evaluative checklists and others for teaching materials checklists. It seems that different disciplines are not benefiting from each other’s research and guidelines. For example, it is clear that educational evaluation depends mostly on applied linguistics, ignoring the contributions of other researchers in other areas like management, business and design. The checklist method is used widely in these fields and many others to evaluate the quality and effectiveness of their programmes and products.

The evaluation checklists are defined as “tools for assessing a product or a service against a set of principles, best practices, or specific criteria” (Brykczynsk 1999, cited in Wilson,

2013: 5). They comprise consulting many references, opinions and constant improvements. Scriven (2007: 5) introduces some of the requirements that should be considered by the evaluator to produce a good checklist where it “should refer to criteria and not mere indicators” as well as being “complete (no significant omissions) and with “contiguous” items that do not overlap and “commensurable, clear, concise and confirmable criteria.” He then divides checklists into six types according to the purpose of their use:

- 1) Laundry list, a set of categories (shirts ... shorts, etc.) that is almost entirely a mnemonic device and very useful just for that reason. In this list, the order of items is not important, “but the entry of entities into the right category on the list is crucial...as well as “the grouping of items.”
- 2) *Sequential checklist*, where the order does matter...One example of this is the preflight checklist, whose use is compulsory...The use of the preflight checklist is evaluative because it is designed to provide support for the evaluative conclusion that the plane is (in certain crucial respects) in good enough condition to fly safely.
- 3) A *weakly sequential checklist* is one where the order is of some importance, but for psychological or efficiency reasons rather than from logical or physical necessity.
- 4) An *iterative checklist* is sequential, in whole or part, but requires—or may require—multiple passes in order to reach a stable reading on each checkpoint.
- 5) Another type of checklist, one that is sometimes but not always sequential, is based on flowcharts. This is the *diagnostic checklist* that is used—for example—by taxonomists, mechanics, and toxicologists.
- 6) Probably the most important kind of checklist for evaluation purposes is the *criteria of merit checklist* (hence, COMlist or, here, comlist). Comlists are widely used as the basis for a particular scoring procedure.

Another classification of checklists is done by Wingate (2002) where there are three main categories: Evaluation planning & management checklists, Meta evaluation checklists and criteria of merit checklists. The relationships between the checklists are explained in two overlapping circles, as figure (1) shows, with internal flexibility of movement of items from one category to another “depending on the nature of the checklist” (Wingate 2002: 2). Checklists that can serve both as evaluation guides as well as criteria of merit checklists are the ones that can be effectively used in teaching materials evaluation.

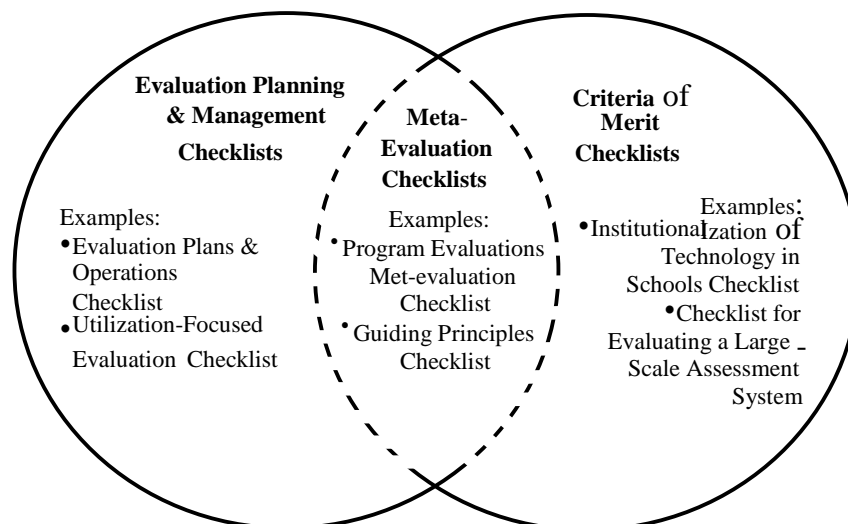


Figure (1) Checklist Categories and their Relationships (Wingate, 2002)

The checklist can be used in three different ways regarding teaching materials: to “borrow and adapt, to originate (to brainstorm ideas) and to research (find out what end users: teachers and students considered important)” (McGrath 2002: 41). The last two are the most difficult as they require plenty of time and thoroughness. Like Scriven and Tomlinson, other researchers (Bicheleymer 2003, Stufflebeam, 2000) produced their own checklists on how to create checklists, which can be also a very useful guide in the development of new checklists. Bichelmeyer (2003) gives very detailed instructions with five main categories: context, content, structure, images and usability. Each one of these main categories is divided into detailed items to be considered when designing a checklist. These checklists can be adapted to use in certain contexts, but they should be simplified and tested to be used with all types of users. Another checklist for developing checklists is provided by Stufflebeam (2000), where his main categories are twelve instead of five. His instructions for developing checklists are very detailed and thorough. Tomlinson (2013) provides the principles that should be followed in developing evaluation checklists for English language teaching materials. But they are too theoretical and difficult to apply in real settings. His principles that are used for evaluation are “derived from Second Language Acquisition and learning theory” where they form the basis for his offered evaluation criteria in (Tomlinson, 2013 a). Some of his principles are easy to detect in teaching materials, but others such as teaching materials “should not expect effective production immediately” are difficult to evaluate unless formative evaluation is conducted throughout the whole course or programme. Tomlinson also encourages the evaluator to use his/ her theory of learning and teaching especially if they

are teachers and they have experience in the field of education. Such views may incorporate the evaluator's opinion and beliefs regarding the students, the teachers and the educational context in general. Tomlinson's recommendations are useful when the evaluators are subject matter experts or trained to conduct systematic evaluations. But for teachers and practitioners, it may be difficult for them to create their own theories of evaluation. Besides, a good evaluation instrument should be based on the experiences of several experts and the testing and feedback of the potential users not just a single researcher or evaluator.

Tomlinson's (2013a) principles for developing criteria or checklists for materials evaluation are included in eleven stages that the developers are supposed to follow in the checklist designing process. In spite of their importance, these guidelines are not conveyed through a tangible or clear basis and sources of the checklists items and main categories. Furthermore, the developing process itself is not illustrated in a conceptual framework, model or any other graphic representation that can simplify its initial different design stages and how it can be used when finished. The heuristics developed by Tomlinson can be very useful, at the same time confusing, for example some items in the middle (5-9). Item (1) also limits the bases of the checklist to two sources; "principles of language learning" and "classroom observations" (Tomlinson 2013 a: 37-44) which make the checklist sources incomplete.

- 1) Brainstorm a list of universal criteria based on principle of language learning and classroom observations
- 2) Subdivide some of the criteria
- 3) Monitor and revise the list of universal criteria
- 4) Categorize the list
- 5) Develop media-specific criteria
- 6) Develop content-specific criteria
- 7) Develop age-specific criteria
- 8) Develop local criteria
- 9) Develop other criteria; teacher-specific, administrator-specific, gender-specific, culture-specific or L1-specific
- 10) Trial the criteria
- 11) Conduct the evaluation

Tomlinson and Scriven's criteria can be used as "assessment rubrics" for the final textbooks evaluation checklists. But they still remain unclear for developers and evaluators, especially novice evaluators, as there are some procedures that need further

explanations. These clarifications can be presented either within the developed checklist or in a separate booklet or a complemented website.

Wilson (2013) also suggests guidelines for developing checklists from different sources (Brykczynski 1999; Burian 2004/2006 and Galib & Graham 1993) to include: “designing the physical form of the checklist...keeping it short but not too short...not too general...not too specific”. The checklist developer has also to make sure that all the “terminology...will be understood by the potential users”, “the phrase items” and “layout and style” formatting are consistent (Wilson 2013: 22-23). Similarly, the questions that materials evaluation checklists developers should pay attention to when finishing the checklist design are listed by Tomlinson (2017: 63) below:

- 1) Is the list based on a coherent set of principles of language learning?
- 2) Are all the criteria actually evaluation criteria or are they criteria for analysis?
- 3) Are the criteria sufficient to help the evaluator to reach useful conclusion?
- 4) Are the criteria organized systemically (e.g. into categories, and subcategories...)?
- 5) Are the criteria sufficiently neutral to allow evaluators with different ideologies to make use of them?
- 6) Is the list sufficiently flexible to allow it to be made use of by different evaluators in different circumstances?

All the above guidelines will help developers who are experienced checklist designers, but they will remain vague concepts to the ordinary practitioners. All the instructions will be almost useless unless they are tested practically with users in real settings. Once more, the guidelines they suggested lack the empirical part that can help in the full comprehension of the phases of the design, the processes and the usage. Validating the developed checklists is as important as their development. When the development explanations are based on the designer's own experience and the users' recommendations are not considered through the authentic use of the developed checklist, any heuristics or guidelines will be understood only by their developers.

The question, then, is how can a viable checklist be developed and used by any practitioner without having to go through the ELT complexities and dilemmas. McDonough et.al (2013: 39) think that “a more straightforward way of looking at this kind of multiple-component syllabus is to see it in terms of merging of two broad approaches. One of these is concerned with a view of language in use, and includes

categories of function, context and language skills. The other is a version of a more formal linguistic syllabus, which comprises elements of grammar, pronunciation and vocabulary". Sheldon (1987:7) suggests that "textbooks and materials need to be evaluated with reference to linguistic theory" and by many stakeholders as they are "evaluated not only by teachers and reviewers, but also by educational administrators charged with obtaining the best text- book value for money". Lim and Lee (2007: 67) mention such impacts that are required for materials evaluation like "curriculum theories, instructional design theories... learning theories and second language learning (SLL) theories." Models that are based on language theory continuum from behaviorism to constructivism include: "cognitive model, structural model, functional models, interactional models, sociocultural model, genre model and lexical model" (Richards & Rodgers 2014: 20-25). These models are the basis for many teaching methods and syllabuses. All of these contributed to materials development and consequently to their evaluation. In their attempt to deal with materials evaluation, Byrd & Schuemann (2014: 381) tried to develop criteria for "textbook evaluation and selection", trying to specify the "conceptual underpinnings" of their scheme. Their foundations of developing the checklist were (1) "theories about the concept of the textbook and the purposes and uses of published sets of materials". An example for this are questions about textbook definition and their roles. The second basis is (2) "theories about the nature of language" and "the nature of language" here is the "linguistics directly or indirectly embodied in a textbook. The third underpinning is (3) "the relationship between textbooks and language teaching/ language learning". Within this area the questions about how the "ideas" of "language teaching/learning are reflected in the textbook" and about the roles of both teachers and students are raised. Though this could be one of the rare attempts to identify the evaluation checklist foundations, the developers did not delineate their conceptual underpinnings through a model or a framework. Therefore, going through their criteria may become difficult for some potential users and developers of teaching materials evaluation instruments. As teachers or beginner developers, having to embrace both the theoretical underpinnings of materials development and evaluation, as well as the tangible textbooks as products with certain characteristics, is very difficult if it is not guided by practical training or a clear evaluation tool that is accompanied by a clear conceptual framework and simple instructions. Hutchinson (1987: 37) indicates how some descriptions on textbooks may become meaningless for the users and consequently difficult to be evaluated and assessed:

Unfortunately it is not always possible to gather such information from the publishers' own descriptions of the materials. What does it actually mean when a course is described as 'communicative' or 'functional'? Is it an accurate description or merely a fashion label? Does the concept of 'communicative' accord with your own interpretation? The terminology has become so meaningless that we cannot rely on it.

Litz (2005: 12) suggests that “one of the most useful starting points in any textbook evaluation is an analysis of the authors' and publisher's credentials”. This in fact can be a good start, but not professional when evaluation is thought of as a process that requires both theoretical and pedagogic backgrounds as “the decisions made by coursebook writers are inevitably influenced by theoretical statements and research outcomes in applied linguistics” (Nunan 1991: 214). In other words, teachers may not be able to understand some of the authors' claims, and consequently they may fail to detect the ‘made-up’ statements about the content of the textbooks. Moreover, the reliance on mere content analysis of the textbooks will not help in making sound decisions about their selection or effectiveness. Harwood (2010) refers to the limitations of content analysis as “quantitative and qualitative forms of content analysis have been critiqued...the former is guilty of enabling breadth at the expense of depth” and the latter “suffers from reliability issues...unless conducted rigorously and systematically” (Harwood, 2010: 11). In addition, some textbooks may deceive the evaluators through their colourful layout and design. Other textbooks may fail to appeal to the teachers and learners who are using them though they may have excellent content. Tomlinson (2013: 4) noted that some textbooks “often lack energy and imagination” and are “sometimes insufficiently relevant and appealing to the actual learners who use them”. Despite that, the teachers and learners are forced most of the time to use such textbooks because they cannot express their dissatisfaction with the selected materials to the programme directors or the authorities as they do not have strong justifications or evidence of their opinions. Being trained in materials evaluation and provided with viable instruments, teachers can be more constructive in the materials use, their evaluation and supplementation. In fact, “all teachers need frequent stimulus and refreshment if they are not to ‘fossilize’ though “most teachers have very few opportunities for personal and professional development” (Tomlinson 2013 a: 9). Therefore, the involvement in materials selection and evaluation is one of the most useful tasks that can help teachers, learners and institutions develop and improve their courses and programmes.

As the textbooks are skillfully marketed, teachers and potential users in general have to be equipped with suitable tools for selection and evaluation. According to Tomlinson, the writers are not the only people to be blamed for the dissatisfaction of the materials, but also “publishers, teachers, institutions and ministries” which may cause “negative effects on learners’ potential to benefit from their courses.” To avoid such problems in the English Foundation Programmes, an established framework for materials evaluation has to be part of the whole programme. Clement (1942: 10) demonstrates that:

The chief object of the use of the proposed general analysis outline or checklist is to guarantee a ready and intelligent acquaintance on the part of all persons involved with what actually exists in textbooks. There is always a great temptation for individuals to begin at once, when examining textbooks, by passing hasty judgements [on their worth].

Kurniawan (2006: 3) noticed that some teachers “undertake evaluation reluctantly since they are not sure what to base their judgments on, or how to qualify these judgements. Indeed they find themselves lacking an appropriate approach to evaluation as literature on this subject is sparse.” In other words, there are many instruments and tools for materials evaluation in the literature, but in practice none of these instruments is applicable to practitioners as most of their approaches and bases are ambiguous to the potential users. As a result, the first step in developing the teaching materials evaluation checklist in this study is to specify its sources and the foundations underlying its design and development.

2.11 The Initial Concept of the Checklist and its Sources

As Ansary & Babaii (2002:2) have noticed, “during the last three decades” textbooks selections and adoption “have essentially been based on ad hoc textbook evaluation checklists.” As a result, “the shaky theoretical basis of such checklists and the subjectivity of judgements have often been a source of disappointment”. To prevent such faults, all possible sources of developing teaching materials evaluation checklist for this project have been considered. Based on the literature review of all the related fields (Second Language Acquisition, design, evaluation, teaching materials, instructional design models, curriculum studies, learning and teaching theories and checklists development), it has become obvious that two main strands underlie the development of any evaluation checklist. They are the theoretical aspect (research) and the contextual aspect (setting needs). Thus, both aspects are to be used to develop a conceptual framework to guide the

design process of the detailed checklist construction and to guide other developers. It establishes a strong theoretical and practical ground for the design of the evaluation checklist as it forms the foundations that can be used as a starting point for any future checklist development for the ELT materials selection and evaluation. This framework will enable a replication of the study procedures and consequently the checklist development. So, replicability becomes possible in design-based research studies in the general processes when the developed instrument is based on a specific and clear framework regardless of the final findings that may differ according the various contexts. To validate the conceptual framework before using it for the checklist design, it was tested with four subject matter experts who have been involved with materials development and evaluation for many years. Following this, prototyping was used to demonstrate the main heading and sub-headings of the initial checklist. The first prototype was designed by the researcher and was based on the analyses emerging from the literature review and included three main categories: SLA principles, ELT curriculum design principles and teaching principles. The second prototype, based on the setting and stakeholders needs, was also developed with three main categories: learners' needs, teachers' needs and institutional' needs. Subsequently, both prototypes were merged into a cohesive prototype for the teaching materials evaluation checklist and the final prototype of the checklist was tested for usability and effectiveness issues through the phases of design-based research methodology.

In the field of materials evaluation, some researchers have been disappointed by the lack of spread and use of evaluation checklists and others have been trying to find alternatives. Tomlinson (2012: 149) states that “in the last ten years a number of... writers have proposed frameworks for materials evaluation instead of checklists”. Those writers are McGrath (2002), McDonough & Shaw (2003), Riazi (2003), Cunningsworth (1995), Wallace (1998), Rubdy (2003), Tomlinson & Masuhara (2004) and McCullagh (2010). Despite their valiant attempts, the frameworks they have presented lack practicality aspects especially for the ordinary teachers in the English language programmes. They are also in the form of descriptive processes that require the evaluator to be familiar with specific terms and to have enough academic background before they can attempt using them. There have been no real endeavors for conceptualization or simplification for novice teachers. As a result, such frameworks have rarely been used for materials evaluation. As Hutchinson (1987) point to: “materials evaluation is essentially a

matching process in which the needs and assumptions of a particular teaching-learning context are matched to available solutions". He believes that "this matching process has four stages. The first is to "define the criteria on which the evaluation will be based" the second is to "analyse the nature and underlying principles of the particular teaching/learning situation", the third to "analyse the nature and underlying principles of the available materials" and finally "compare the findings of the two analyses" (Hutchinson 1987:41). Moreover, Brown (1995: 232) specified that the sources of "sound evaluation practices will be based on all available perspectives so that many types of information can be gathered to strengthen the evaluation process and ensure that the resulting decisions will be informed, accurate, and useful as possible." In relation to the setting needs, Ansary & Babaii (2002) stated that "teachers, students, and administrators are all consumers of textbooks" and "may have conflicting views about what a good/standard textbook is." Their question was "where they can turn to for reliable advice on how to make an informed decision and select a suitable textbook" as the "literature on textbook selection and/or textbook evaluation procedure is vast" (Ansary & Babaii 2002: 3). It is not only vast, but confusing and ambiguous for most stakeholders. With regard to research, Second Language Acquisition along with teaching approaches and curriculum design principles are the most important factors to consider when attempting to develop or use teaching material evaluation instruments. They integrate and contain the answers for the questions about the content of materials (what is there), the pedagogical factors or the teaching principles that guide the teaching processes (how to deliver content) and the ways both content and its delivery are organized (curriculum design). The previous frameworks and criteria lack comprehensiveness because they fail to focus equally on both the theoretical aspects and the contextual aspects. Most criteria also fail to include "a combination of both approaches" of the evaluation: the "impressionistic overview of the whole and the in-depth evaluation" which "will make a sound basis for evaluation and for ensuring choice of the most suitable textbook for adoption" (Kurniawan 2006: 4).

Byrd & Schuemann (2014) based their evaluation on the "textbook fit with the curriculum, the "textbook fit with the students" and the "textbook fit with the teachers". Though they provided general questions to be asked for each aspect, their framework lacked clarity and inclusiveness. They ignored other stakeholders involved in evaluation, such as authorities, and other theoretical aspects in research, where they focused only on the materials fit with the curriculum. Though they thought that both the categories of

textbooks selection before use and textbooks evaluation while use are the same, they considered that they differ in “the purpose”, so their checklist is basically based on the same categories for both selection and evaluation. Ansary & Babaii (2002: 5), attempted to select “a set of common-core summary characteristics” that “can be identified as universal” and which appeared to them “across the reviews” of ten textbooks evaluation studies and the main categories in ten popular checklists. Their aim as they stated is “at the very least, probably the application of a set of universal characteristics of EFL/ESL textbooks may well help make textbook evaluation a coherent, systematic and thoughtful activity” (Ansary & Babaii 2002: 6). Again, the purpose is very important, but the method is not comprehensive and lacks lucidity. The two main constituents of any evaluation tool are research findings and setting needs, not only previous experiences of others. They are the starting point for any attempt to design a teaching materials evaluation tool. Three fixed bases should be considered in research: SLA, teaching principles and research on ELT curriculum. Another three bases on setting should be involved to include the main stakeholders: students, teachers and institution. A summary of Tucker’s (1975: 359-361 cited in Ansary & Babaii 2002: 7-8) characteristics of a good evaluation criteria include:

- 1) a predetermined data-driven theory-neutral collection of universal characteristics of EFL/ESL textbook, discrete and precise enough to help define one's preferred situation-specific criteria,
- 2) a system within which one may ensure objective, quantified assessment,
- 3) a rating method that can provide the possibility for a comparative analysis,
- 4) a simple procedure for recording and reporting the evaluator's opinion,
- 5) a mechanism by which the universal scheme may be adapted and/or weighted to suit the particular requirements of any teaching situation,
- 6) a rating trajectory that makes possible a quick and easy display of the judgments on each and every criterion, and
- 7) a graphic representation to provide a visual comparison between the evaluator's preferred choices as an archetype and their actual realizations in a particular textbook under scrutiny.

These characteristics are important to consider before attempting to develop an evaluation instrument for teaching materials. Most of the available criteria focus only on a few items of the above list, which results in producing incomplete schemes for the evaluation of teaching materials in the English language programmes.

2.12 Summary

In order to develop or use teaching materials evaluation instruments any practitioner, evaluator or developer will come across difficulties on several levels such as the interdisciplinary area of materials development and evaluation, the various methods used in evaluation and lack of clear model to follow in developing teaching materials evaluation instruments. The suggested passage, through this study, for a successful journey to a comprehensive understanding of teaching materials evaluation instruments development will incorporate several stages. First, developers or evaluators should be acquainted with the meaning of evaluation in general and its models and theories. Second they are advised to have sufficient background about curriculum and materials development. They have to be familiar also with the role these materials play in English language programmes, to be able to develop a practical and reliable set of evaluation criteria. Finally, teaching materials evaluation checklist developers have to recognize the different methods of teaching materials evaluation which will enable them to base their instrument on clear foundations and bases. Unfortunately, the previous evaluation instruments fail to familiarize the evaluation developers and users with such issues including the most successful tools for teaching material evaluation, the checklists. This can be attributed to several reasons such as depending on previously-developed instruments and consequently copying most of their features, or basing the developed evaluation instruments on one source such as Second Language Acquisitions principles, or failing to identify the real sources of their evaluation tools and accordingly the conceptual framework that demonstrates their developments. Or maybe just failure to use the appropriate methodology (e.g. design-based research) to construct and design their evaluation tools, a methodology that will enable them to connect all of the mentioned issues within one study as chapter three will exemplify.

Chapter 3: Research Design and Methodology

3.1 Introduction

Chapter two exemplified the problems and methods of teaching materials evaluation in the English Foundation Programmes in Colleges of Applied Sciences in Oman. In these programmes, developing evaluation instruments for the selection and the evaluation of teaching materials before use, while-use and after use is an essential part for their improvement. Likewise, finding the methodology that can incorporate design, theory and practice is an important aspect for the success of the instrument design and development. In this study, design-based research methodology is thought to be the appropriate method for developing such evaluation instruments. Shah et.al (2015: 152) referred to “the role that design-based research (DBR) can play in addressing the complexity of education” through “informing immediate practice while simultaneously continuing to develop theoretical understandings in the field of education.” It allows researchers to study the topic from different angles which result in a comprehensive treatment and understanding of the teaching materials evaluation in particular and other related fields in general. The DBR iterative and pragmatic nature can empower the researchers to move between the macro or holistic approaches and the micro approaches of conducting research that seek “scientific reasoning in atomic building blocks of human action” (Goldkuhl, 2004: 17) in clear and practical methods. In this study DBR aims to connect the various processes of the evaluation instrument development which comprise (i) constructing a conceptual framework, (ii) testing it with experts and users, (iii) using it to design the teaching materials evaluation instrument and consequently (iv) reviewing the final prototype of the evaluation checklist with different stakeholders. Collins et al., (2004: 15) expound that “design experiments were developed as a way to carry out formative research to test and refine educational designs based on theoretical principles derived from prior research” which involves “putting a first version of a design into the world to see how it works” and after that the “design is constantly revised based on experience, until all the bugs are worked out.” The design-based research phases are exemplified through this study. The first phase (exploration and analysis) in DBR which included many research activities aimed at understanding both the research topic through literature and the local context of the study through needs assessment. Based on the literature reviews within this phase, the general theoretical sources for the teaching materials evaluation checklist

development were specified. These underpinnings were Second Language Acquisition principles, the teaching principles and the ELT curriculum. Then the setting needs were investigated through brainwriting sessions and a short survey. Afterwards, these foundations are illustrated through a conceptual framework that is developed by the researcher. Then, the conceptual framework for the teaching materials evaluation checklist is tested with subject matter experts which eventually led to the checklist propositions, categories and items. The conceptual framework, depicting the checklist sources and main checkpoints or categories, which was one of the activities in phase one is the starting point for the checklist construction (phase two). The processes of the conceptual framework are clarified next before delineating the position, definitions and models of design-based research.

3.2 A Conceptual Framework for the Evaluation Checklist Development

Designing an evaluation checklist is a multi-tasking job where the developer has to connect various theoretical and empirical tasks to yield one coherent product. Therefore, basing checklist design on just one source (e.g. Second Language Acquisition as some developers do) ignores the sophisticated process of materials design and evaluation. Based on literature and previous studies, it has become clear that two main aspects need to be consulted in developing English language teaching materials and their evaluation: research and setting needs. As Lim & Lee (2007) elucidated the theoretical and the practical considerations and the pedagogical concerns are all important in designing teaching material checklists which may include “various curriculum theories, instructional design theories and learning theories” Lim & Lee (2007: 67). A robust checklist needs to be based on solid theoretical assumptions from research and on a consideration of the setting where it will be used, in order to ensure that the checklist fulfills its potential users’ needs. The literature on the hypothetical strands that contribute to materials evaluation instrument in this study include three main sources: Second Language Acquisition principles, teaching approaches and the ELT curriculum. The practical considerations related to the setting needs (students, teachers, coordinators and authorities) are also included as another source for evaluation instrument development. These sources are connected together through a conceptual framework and the detailed description of the framework development and validation is demonstrated in chapter four. The established conceptual framework assisted in the development of the teaching materials evaluation checklist with its different design processes as well as its review and

validation cycles within the various phases and iterations of design-based research methodology. For a better understanding of design-based research methodology and its position in relation to other research paradigms, a short introduction is provided next.

3.3 DBR within Common Research Paradigms

3.3.1. Introduction

The word paradigm was “first termed by Thomas Kuhn in his 1972 book, titled *the structure of Scientific Revolutions*, and it “refers to an overall theoretical research framework” (Mack 2010: 5). Another definition is introduced by Willis suggested that a paradigm is “a comprehensive belief system, world view, or framework that guides research and practice in a field” (Willis 2007 cited in Taylor & Medina 2013: 1). From a philosophical perspective, “a paradigm comprises a view of the nature of reality (i.e., ontology) – whether it is external or internal to the knower; a related view of the type of knowledge that can be generated and standards for justifying it (i.e., epistemology); and a disciplined approach to generating that knowledge (i.e., methodology).” It is also defined by Creswell & Plano Clark (2007 cited in Hall, 2013: 3) as “worldview”. The “the four commonly agreed on worldviews” as stated by those researchers “are post-positivism, constructivism, transformative and pragmatism” (ibid: 3). These paradigms differ in their ontological, epistemological and methodological assumptions based on their views of reality and how to approach it. Generated from these paradigms and others “positivism, realism, constructivism and critical theory” (Sobh & Perry 2006: 1195), the two main theoretical approaches or perspectives (quantitative and qualitative) are used as research methodologies to guide researchers in their quest for truth. Mixed method introduced later as a third methodology that combines both quantitative and qualitative. Mixed method appeared as a result of so called “paradigm wars of the 1970s and 80s where the positivist paradigm of the quantitative research came under attack from social scientists supporting qualitative research and proposing constructivism” (Hall 2013: 2). Some new approaches can easily be located within the appropriate paradigm especially if they can be found at the far end of the paradigm continuum: positivism and constructivism. For example, “multi-methods do not have the same paradigmatic problem as do mixed method since they can adopt the paradigm appropriate to the single type of data being collected” (Hall 2013: 2). A researcher using a multi-method methodology can use two qualitative or quantitative methods in the same study without having to worry about the study’s theoretical framework. But, other methodologies,

especially mixed methods and new methods such as design-based research, may require more time and effort from the researchers. According to (Walker 2001: 53) “a very noticeable aspect of the design research literature is the absence of discussion of epistemological issues” and despite the fact that “the word is not totally absent from the general DBR literature, there is no serious discussion of epistemological issues.” This can be attributed to the confusion in positioning such methodologies in the research paradigms continuum. So, clarifying its position will help to avoid such conflicts and simplify its use in educational research studies.

3.3.2 DBR in a Critical Realism Paradigm

In searching for the position of Design-Based Research within the paradigms outlined above, some researchers have suggested that “Critical realism (CR), largely based on the writings of Bhaskar (1975, 1978, 1989) is the appropriate paradigm as it is “often seen as a middle way between empiricism and positivism on the one hand, and anti-naturalism or interpretivism on the other” (Zachariadis et.al 2013: 858). This view is explained by Andriessen (2008: 126) in detail:

Advocates of design-based research share an epistemology rooted in pragmatism (Romme, 2003; Wicks & Freeman, 1998). However, they may differ in their ontological point of view. I believe in the ontology of embodied realism (Lakoff & Johnson, 1999) but alternative positions may include critical realism, historical realism, and relativism (Lincoln & Guba, 2000). In addition, Van Aken and Romme (2005) argue that researchers can draw from several different research methods to test the validity of the design, ranging from more positivistic quasi-experiments (Cook, 1983) to action research type interventions (Susman & Evered, 1978). This implies that design-based research may make use of a variety of methodologies.

It is clear that there is a mix between the concepts of ontology and epistemology in relation to design-based research. It is true that design based research can make use of a variety of epistemologies, but that does not mean that it can belong also to more than one paradigm. Design-based research may share some concepts with a critical paradigm such as “raising the conscious awareness” and making a “difference” (Taylor & Medina, 2013: 3) for participants and stakeholders, but DBR does that through being a revealing methodology rather than a critical one. Thus, change in design-based research can be achieved through different lenses other than criticality, which are understanding and recognition of the real contexts difficulties and the genuine determination to solve their

problems and create innovative and sustainable solutions that can make huge differences. In fact, one can say that an approach like design-based research where it incorporates theoretical abstractions, practical design and evaluation or reviewing activities is difficult to align with specific paradigm. But as its main purpose is to produce useful and practical products and innovations, Goldkuhl (2011) conception about design research can be applied here. He stated that “there is a spectrum of different, but related epistemic types in design research” and “this “epistemic spectrum corresponds better with pragmatism than positivism, interpretivism or critical realism” Goldkuhl (2011: 89). The spectrum of the epistemic types specified by Goldkuhl are “related to four aspects/ types of pragmatism: local functional pragmatism (as the design of a useful artefact), general functional pragmatism (as creating design theories and methods aimed for general practice), referential pragmatism (focusing artefact affordances and actions) and methodological pragmatism (knowledge development through making) Goldkuhl (2011: 84). In this study, design-based research, with its pragmatic assumptions, is taken to be a ‘one size fits all’ alternative for educational problems. Hence, to simplify its use and to avoid such complex views about its philosophical foundations, it is recommended to include it under the umbrella of the pragmatism paradigm as the “majority of writers in the literature on research methodologies agree that pragmatism is the paradigm that can be the underpinning philosophy for design-based research” (Barab & Squire 2004, Juuti & Lavonen 2006, cited in Alghamdi & Li, 2013: 2).

3.3.3. DBR in a Pragmatism Paradigm

The Pragmatism paradigm “was first introduced through the works of Charles Sanders Peirce (1839–1914), and then further developed by William James (1842–1910), and John Dewey (1859–1952)” (Given 2008 cited in Alghamdi & Li 2013: 2). It does not follow any specific methodology, rather this paradigm puts “the research problem as central and applies all approaches to understanding the problem” (Creswell 2003: 11 cited in Mackenzie & Knipe, 2006: 196) .As a result, researchers can use any method that suits their research design, whether qualitative, quantitative, mixed methods, multi-method or other methods like evaluation techniques.

Hogue (2013: 1916) postulate that “educational design research draws influence from the design sciences, such as architecture and engineering... in addition to its pragmatic underpinning”. Collins, Joseph, & Bielaczyc also considered design-based research as a “pragmatic approach” in their definition, which suggests that “education design research

is a pragmatic approach to research with the dual goals of (1) solving an educational design problem in a real-world context, and (2) contributing to scholarly knowledge in the form of instructional design theory or design principles” (Collins et. al. 2004: 19).

Another method that is confused with DBR is mixed methods. As MacKellar (2010: 20) explains “design-based researchers are not alone in no man's land... these two groups share more than the terrain between two entrenched opponents; they also share a desire to avoid philosophical posturing and debates...pragmatically, they have chosen to take themselves out of the melee so that they can get on with the work of research”. In the literature, mixed methods are not considered as a robust methodology by some scholars compared to quantitative and qualitative methodologies. Despite that, mixed methods can play an important role in educational research and instead of calling for the “death of mixed methods” (Symonds & Gorard 2008: 15), attention can be shifted on how to exploit their use to complement other methods in conducting research such as DBR studies. The attempts to recommend mixed methods as a third methodology is causing more confusion and uncertainties on the level of paradigms (realism, constructionism, pragmatism) and the level of methods (data collection instruments, analysis and sampling). Attempts to create certain sampling techniques is another problem that faces researchers using this “method” as they have also to think about their data analysis, and their integration at certain points of the research project. Historically, the cause of the origination of this method (during paradigms war period) is to utilize both quantitative and qualitative data collection methods in the same study. Akilli (2008: 6) discussed the differences between “design-based research (DBR) and mixed methods research” and though there are some similarities in utilizing a pragmatic approach, they “are two different entities.” As mixed methods main contribution is using both quantitative and qualitative methods, the contribution of design based research is making use of both theory and practice to solve educational problems. In fact design-based research according to Akilli, (2008: 6) can offer:

a new worldview of theory development and refinement along with design to construct design sciences of education... it also offers a newly-emerging research methodology, which draws from different fields of design and education and even includes mixed methods approach... It offers usable knowledge that informs theories which in turn inform real-world practices. It produces dynamic knowledge that changes dynamically in relation to context, which is a dynamic, winding structure that is shaped by time, place (space), actors and actions that take place.

It offers, local knowledge, since it produces tentative generalizations that are drawn from initial implementations, which makes DBR a local science.

Design-based research can not only utilize mixed method approach, but it can borrow and embrace methods and techniques from other different types of research. Though they may share some perceptions, design-based research's distinctiveness can be realized when it is compared to these types of research.

3.4 Design-based Research and other Types of Research

Characteristics of design based-research are compared to several types of research in terms of similarities and differences. For example, Reimann (2011) mentions the types of research where design-based differs on certain levels despite sharing some similarities. He indicates that “DBR is different from curricular studies (DBR has more of a focus on the enacted curriculum), evaluation studies (focus on process, not only outcomes), pure discourse studies of classroom talk (multiple methods, including quantitative ones, are used), action research (aspiration to theory development), lesson studies (not confined to learning in classrooms and through teachers), and instructional design (learning theory development)” Reimann (2011: 39). In the next sections, DBR is discussed in comparison with design research, action research and evaluation research.

3.4.1 DBR and Design

As its name suggests, design is a major part of Design-Based Research, which may include designing any innovative tools or instructions such as programmes, products, processes or policies mentioned by McKenney and Reeves (2012) or any other educational instruments. This corresponds to Friedman's (2003: 507) description of design which “refers to a process that involves creating something new (or reshaping something that exists) for a purpose, to meet a need, to solve a problem or to transform a less desirable situation to a preferred situation.” Though it is a “highly systematic, problem solving process, executed by individuals who specialize in portions of the larger process, and informed by empirical evidence gathered throughout the design process” (Smith & Boling 2009:13), design processes alone cannot be considered a rigorous research methodology. Additionally, “the skills of creative designers and the attributes of rigorous scholars have limited overlap...when designers receive formative feedback, their intuition often leads to changes that may neither be grounded in theory nor be limited

to enable comparative research across time” (Dede, 2005: 6). Designers’ ultimate goal is only to produce practical knowledge rather than theoretical understanding.

DBR goes far beyond simple design and seeks to integrate the advantages of both design and research. Shavelson et.al., (2003 cited in Dede, 2004: 235) propose that design research is “based strongly on prior research and theory and carried out in educational settings” and it “seeks to trace the evolution of learning in complex, messy classrooms and schools” to “test and build theories of teaching and learning” in order to “produce instructional tools that survive the challenges of everyday practice” .Therefore, “this definition implies three important, deeply intertwined goals for design-based research projects—research, design, and pedagogical practice” (ibid: 235). Four main trajectories can be found within a design-based research study, which include the “product usability trajectory (implementation and effect), the “product design trajectory (changes and amendments) and “the research trajectory (reporting usability effects and amendments) and finally “building design theory and heuristics). In contrast, design studies can make use of only the first two trajectories to accomplish their purposes and aims neglecting the theoretical aspects and design guidelines and heuristics.

Unlike simple design studies, design-based research can be placed among other types of research considering Stokes' (1997) quadrants diagram. This diagram divides research into three main quadrants in terms of their final purposes (whether the main purpose is to produce theoretical knowledge, practical knowledge or a combination of both). The three quadrants are the Bohr quadrant which represents pure basic research, the Edison quadrant representing applied research and the Pasteur Quadrant with the name ‘user-inspired research’ representing research which seeks to produce both theoretical and practical knowledge. This classification places design-based research “directly at the center of Pasteur’s quadrant” according to Bannan-Ritland (2003) and Roschelle et.al. (2011) reported in Shah et al. (2015: 156) which means that design-based research aims to produce theoretical and practical knowledge. This is clarified by Edelson (2002: 112), when he explains that “the goal of ordinary design is to use the lessons embodied in a design procedure, problem analysis, and design solution to create a successful design products. Design-based research retains that goal but adds an additional one, the goal of developing useful, generalizable theories about the developed instruments.” Design-based research is also different from action research which will be explored next.

3.4.2 DBR and Action Research

Practitioner research is defined by Richardson (1994 cited in Anderson & Herr 1999), as “practical inquiry” which focuses on the “improvement of practice”; she then uses her own definition “to relegate it to secondary status vis-a-vis formal (“real”) research...” This definition and debate is true regarding design-based research and “clearly the formal/practical knowledge debate is about more than research epistemology and methodology; it is about the very nature of educational practice itself.” (Anderson & Herr 1999: 15). Design-based research is an empirical research, but not in laboratory nor control groups, rather it is carried out in the real places of the intervention or the experiment. It incorporates both the iterative cycles that focus on the product itself or intervention and research rigor and soundness. Some researchers consider DBR as a type of practitioner research which is an “inquiry by practitioners themselves on educational problems, designed to improve practice and share solutions with others” Wilson (2014: 6). Wilson (2014) includes besides DBR, action research and evaluations studies, which Wilson defines as the “local studies designed to assess progress or impact, or improve programs and services” (Wilson 2014: 7). Design-based research and action research can be similar on methodological level where it is difficult to identify the appropriate methods that are compatible with the methodology used for conducting research studies. Ewing (2011) proposes that “a research methodology should be seen as an inter-related set of philosophical assumptions, rather than a technical process that must fit one set of particular conventions...these assumptions and principles have implications for every step of the research, from the questions identified as appropriate for the investigation, to the nature of the data needed, to the methods that are employed, to the analyses that are appropriate and, finally, to the claims that can be reasonably made or the conclusions that can be drawn” Ewing (2011: 71-72). Thus, “while methods are the tools employed to study a phenomenon” the “methodology applies to the principles underlying them” Groundwater-Smith & Irwin (2011: 57). And in “action research the theoretical analyses are far more eclectic than the term ‘methodology’ suggests” as Groundwater-Smith & Irwin (2011: 58) suggest. Despite the explicit similarity on the level of the methodology, they differ implicitly as it is difficult for action research to be considered as a research methodology that can be used on its own to answer the research questions. On the other hand, design-based research through its phases can form a comprehensive methodology that can be used to guide the research project and to answer the research questions in a consistent and practical way. Moreover, action research “lacks the emphasis on finding

public knowledge that is a hallmark of educational design research” McKenney & Reeves (2014: 134). Actually, Andriessen (2008: 129) used action research in his study to “test a solution concept” he developed. Thus, action research can be used alone in educational research or used as an accompaniment method of a main study methodology whereas design-based research can only be used as the sole main methodology in different research studies.

Vasalou (2015: 3-4) refers to some differences between action research and design-based research as they can differ in “the design focus and basis” and the strength of produced theory as “the theory-building phase and process are more powerful in DBR” than “in AR”. And though DBR is “newer” it is “more increasingly utilised than AR.” Both design-based research and action research depend on the full understanding of the research context. Even if the design-based researchers are outsiders to the setting, they should familiarize themselves with the environment where they are conducting their research. But at the same time and if “they are 'native' to the setting, they must work to see the taken-for-granted aspects of their practice from an outsider’s perspective” (Anderson, Herr, & Nihlen, 1994: 27 cited in Anderson & Herr 1999: 15).

3.4.3 DBR and Evaluation

Evaluation is also another approach that is seeking rigor through alignment with established paradigms. Despite a myriad of studies, it is still considered a method or an instrument rather than a methodology. In the past, evaluation was directly associated with quantitative research, as the purpose of most research is to produce an objective report which describes the positives and negatives of the evaluand in a fair and unbiased process. Therefore, most studies rely on the positivist paradigm in their search for knowledge, which entails the use of its quantitative methods and instruments for data collection. However, the use of quantitative methodology in evaluation has its drawbacks as all paradigms are “inventions of the human mind and hence subject to human error” (Lincoln & Denzin 1994: 108). Hence, the call for more practical and context-oriented methods such as evaluation studies, that represent human nature, has become accepted and preferred. Later, qualitative or illuminative evaluations are considered more informative about educational problems and contexts. However, when it comes to designing a programme that needs constant changes and different users’ involvements, with their different views and beliefs, another method has to be considered. And in this case, design-based research can be the appropriate alternative. As McKenney & Reeves (2014: 134)

point to the “problem definition and solution design are rarely featured in evaluation research”. Also, they mentioned a key difference where “evaluation research is primarily concerned with evaluating and possibly improving the qualities of a particular intervention...and the broader scientific orientation of generating usable knowledge...is not as overtly present in evaluation research as in educational design research” .For example, in this study designing an evaluation tool for teaching materials in the English language programmes requires more than one study phase. So, the triangulation in methods, places, participants, data collection instruments as well as in the prototypes of the designed tool itself allowed the analysis and the exploration of the problem, the designing of the instrument, checking its usability and specifying validation method besides the yielding design guidelines and instructions of use. These different processes cannot be achieved through evaluation studies, but the iterations of designed-based research allow for such multiple activities of the developed innovations and instruments.

3.4.4 Summary

Design-based research can be a comprehensive research methodology that can take advantage of traditional research, design studies, action research and evaluation studies. Each research method has its specific aims and purposes. For example, ethnography studies, according to Collins et.al. (2004) “provides qualitative methods for looking carefully at how a design plays out in practice” and “large-scale studies provide quantitative methods for evaluating the effects of independent variables on the dependent variables” whereas “design experiments are contextualized in educational settings, but with a focus on generalizing from those settings to guide the design process. They simply “fill a niche in the array of experimental methods that is needed to improve educational practices” (Collins, Joseph, and Bielaczyc 2004: 21). As mentioned, design –based research with its pragmatic nature “draws from many types of research” but despite that “there are several aspects that make it uniquely different from basic, action, evaluation, and applied research” (Shah et.al. 2015: 156). These similarities and differences are detailed by Shah, Ensminger, & Thier (2015: 157) for each type of research:

Similar to basic research, DBR studies seek to expand and refine the knowledge of theory... Similar to action research, which employs iterative cycles related to problem identification, solution generation, DBR studies employ an ongoing approach to research that utilizes multiple phases...the de-emphasizing of theoretical knowledge generation in action research differentiates the two forms (Morgan, 2013). Akin to evaluation research, DBR studies generates knowledge

to inform stakeholders about the value of an innovation...additional purpose of generating knowledge about theory differentiates it from evaluation research. Similar to applied research, which aims to test the theoretical ideas and understanding within a natural setting, DBR studies go beyond the testing of theories by engaging in ongoing cycles of study that involve a systematic process of designing, developing, and implementing innovations to directly address education problems in real time.

Besides the above characteristics of design-based research which make it unique and different from other types of research, it is found to be the most appropriate methodology for answering all four questions raised within this study. More purposes for using design-based research are presented in the following section.

3.5 Rationale for Design-Based Research

The emerging of design-based research is attributed to Collins (1992) and Brown (1992) who “both published landmark papers, which have often been credited as primary catalysts for launching the genre of educational design research” (McKenney & Reeves 2012: 11). For example, Collins (1999 cited in Collins et.al. 2004: 20-21) “compared laboratory studies of learning to design experiments in terms of seven contrasting aspects of their methodology”: “laboratory settings vs. messy situations ; a single dependent variable vs. multiple dependent variables; Controlling variables vs. characterizing the situation; fixed procedures vs. flexible design revision; social isolation vs. social interaction; testing hypotheses vs. developing a profile” and finally “experimenter vs. co-participant design and analysis” (Collins,1999 cited in Collins et.al. 2004: 20-21). In reaction to such practical thinking of educational studies, the inquisitiveness in design-based research has increased, as explained by Anderson & Shattuck (2012) who reviewed and analyzed “the five most cited DBR articles from each year of this past decade” and concluded that the “interest in DBR is increasing” (Anderson & Shattuck 2012: 16). Despite the publishing proliferation on the design-based research, studies and projects that are using this methodology are still few compared to other methodologies.

Education is a highly researched sector and yet practical innovations have never been proven through accumulative and iterative studies that can pinpoint the best practices to use in educational settings. The “majority of educational research literature has acknowledged that educational research is often divorced and alienated from our educational issues and daily practices” (Juuti & Lavonen 2006, Sari & Lim 2012 cited in Alghamdi & Li 2013: 3). The lack of impact could be attributed to the predetermined

procedures and methods used by positivism, for example, that leads to expected results which may not depict the real and complicated context of the study. It could be also the result of researchers' failure to disseminate the results of their research, the disinterest of practitioners and teachers or the intricacy of research language that keeps it far from public and different stakeholders use. Besides, most of the methods and instruments of data collection are used for the final results with no chance to test them again or apply them in real settings. Reeves (2011: 2) explains the reasons behind such failures:

Olson (2004) criticized Slavin's (2003a) enthusiasm for modeling educational research on the randomized controlled trials (RCTs) used by medical researchers. Among the problems Olson identified were that double blind experiments are impossible in education, that implementation variance in educational contexts often severely reduces treatment differences, that causal agents are under-specified in education, and that the goals, beliefs, and intentions of students and teachers affect treatments in ways that are often unpredictable.

There have been constant debates between academics and practitioners about the usefulness and practicality of educational research. There is always that gap which prevents not only the effectiveness of educational research, but also the communication between researchers and practitioners. Anderson, Herr, & Nihlen (1994, cited in Anderson & Herr: 1999: 15) explain the causes of such a gap where "academics (outsiders) want to understand what it is like to be an insider without 'going native' and losing the outsider's perspective" and "practitioners (insiders) already know what it is like to be an insider, but because they are 'native' to the setting, they must work to see the taken-for-granted aspects of their practice from an outsider's perspective." Design-based research is the research methodology that can solve not only educational problems but also work as a mediator between those two parties which will eventually facilitate the spread and use of educational research findings in the different educational contexts and among all stakeholders. This type of research can be considered the link that connects research to practice as most educators complain about the little impact of research theories on practice where "theory and research findings often functioned as little more than slogans for reformers" (Walker 2006: 9).

The same affirmation is also made by Cobb and others as they criticize the available research paradigms in education, such as positivism and constructionism, because "general philosophical orientations to educational matters—such as constructivism—are

important to educational practice, but they often fail to provide detailed guidance in organizing instruction” (Cobb, et al. 2003: 10). Contrary to the traditional research methods, design based research enables more than one phase of data collection and analysis and each phase contributes to the betterment of the intervention or the designed instruments. Examples of individual successful studies, as mentioned by McKenney & Reeves (2013), include Barab, Gresalfi & Ingram Goble (2010), Clark & Dede (2009), and Swan (2007). These “individual DBR studies yielded clear improvements in practice” (McKenney & Reeves 2013: 6) which indicate the importance and future of such studies in the development and advances of educational theories and practices.

3.6 Design-Based Research definitions and models

3.6.1 Definitions

Design-based research is one of several terms that are used to refer to the same approach. Examples of these names include “Design research” (Bannan-Ritland 2003, Collins et. al. 2004, Edelson 2002, Kelly 2004, Reeves et. al 2005); “design-based research” (Barab and Squire 2004), Bell 2004, Bereiter 2005, Design-Based Research Collective 2003, Hoadley 2004, Joseph 2004, Sandoval & Bell 2004); “Design experiments” (Brown 1992, Cobb et. al. 2003, Collins et.al. 2004, McCandliss 2003), “formative research” and “formative experiments” (Reigeluth & Frick 1999, Reinking & Bradley 2004), and design “studies” (Shavelson et. al. 2003, Kelly 2004). Developmental research and educational developmental research are also used to refer to design-based research (Conceicao et. al. 2004, Oha & Reeves 2010 as mentioned by McKenney & Reeves 2008). Kopcha, Schmidt & McKenney (2015) in their special issue editorial on educational design research refer to design-based research as “an emerging approach that attempts to bridge the demand for rigorous research with the development of relevant solutions to educational problems” where educational design research “constitutes a family of design-oriented approaches to educational research, including but not limited to design-based research, design and development research, and design-based implementation research” Kopcha, Schmidt & McKenney (2015: i). Also Richey & Klein (2005) differentiate between developmental research and design-based research, but eventually what they call ‘developmental research processes’ are representing design-based research features and procedures. In this study, the term ‘design-based research’ will be used throughout, as it is the most used and popular term among researchers. Besides, it is thought that the

lessening of the terms used for this types of research will reduce the misperceptions and consequently simplify its use in educational research.

Design-based research has been defined as “a process that integrates design and scientific methods to allow researchers to generate useful products and effective theory for solving individual and collective problems of education” (Easterday, et al. 2014: 6). In fact, “Design research in education is directed at developing, testing, implementing, and diffusing innovative practices to move the socially constructed forms of teaching and learning from malfunction to function or from function to excellence” (Kelly, 2003 Cited in Kelly, et al. 2008: 2). In addition, Barab and Squire (2004: 2) state that design-based research is a “series of approaches, with the intent of producing new theories, artefacts, and practices that account for and potentially impact learning and teaching in naturalistic setting.” Further, Wang and Hannafin (2005: 6–7) define it as “a systematic but flexible methodology aimed to improve educational practices through iterative analysis, design, development, and implementation, based on collaboration among researchers and practitioners in real-world settings, and leading to contextually-sensitive design principles and theories”

A recent definition by McKenney & Reeves (2013: 7) describes educational design research as “a genre of research in which the iterative development of solutions to complex educational problems provides the setting for scientific inquiry.” They further clarified that in Spector, et al.(2013) to include “the solutions that result from educational design research can be educational products (e.g. a multiuser virtual world learning game), processes (e.g. a strategy for scaffolding student learning in online courses), programs (e.g. a series of workshops intended to help teachers develop more effective questioning strategies), or policies (e.g. year-round schooling)” (McKenney & Reeves, 2013: 131). In a more detailed definition, it “can be defined as the systematic study of design, development, and evaluation processes with the aim of establishing an empirical basis for the creation of instructional and non-instructional products and tools and new or enhanced models that govern their development” (Richey & Klein 2007: 1). Shavelson, et al. (2003: 25) emphasize that design-based research is “based strongly on prior research and theory and carried out in educational settings, seeks to trace the evolution of learning in complex, messy classrooms and schools, test and build theories of teaching and learning, and produce instructional tools that survive the challenges of everyday practice”. Of the many definitions illustrated above, Richey & Klein (2007) and

McKenney & Reeves (2013) are the most appropriate and inclusive for the goal of this study. Their descriptions do not confine DBR to educational technologies or classroom settings, rather the definitions encompass any research on any “instructional and non-instructional products and tools” (Richey & Klein 2007: 1) that help to enhance the educational field with its complicated “products, processes, programmes and polices” (McKenney & Reeves 2013: 10). These two explanations broaden the approach of DBR, which makes it more applicable to any research project including the development of teaching materials evaluation checklist in this study. In fact “technology development has not been the driving force behind DBR” as “DBR was and is mainly concerned with innovations in teaching and learning that pertain to pedagogy rather than technology” Reimann (2013: 45). Moreover, Design-Based Research Collective (2003: 8) specify “four areas where design-based research methods provide the most promise” to include : (a) exploring possibilities for creating novel learning and teaching environments, (b) developing theories of learning and instruction that are contextually based, (c) advancing and consolidating design knowledge, and (d) increasing our capacity for educational innovation”. Beside all of these areas, it is thought that design-based research is a methodology that can be used in investigating any topic in the educational field as well as other fields and disciplines. If promoted in postgraduate research studies, DBR can be used in solving learning and teaching problems, designing innovative technological interventions and developing many educational products, instruments and programmes.

3.6.2 Models of Design-Based Research

Design-based research characteristics as described by McKenney are interventionist (it aims to improve practice); collaborative (it requires multiple participants and varied expertise); responsively grounded (the findings from one phase influence subsequent directions); and iterative (it anticipates multiple cycles of inquiry and action) McKenney (2013: 5). There are many models that demonstrate the different qualities, phases and cycles of design-based research. Reeves’ (2000) model, revised in 2006, represents four phases in very simple and comprehensible tables as displayed in Figure (2). Several models can be found in the literature, but the most practical ones will be investigated to select the appropriate one for this study’s processes.

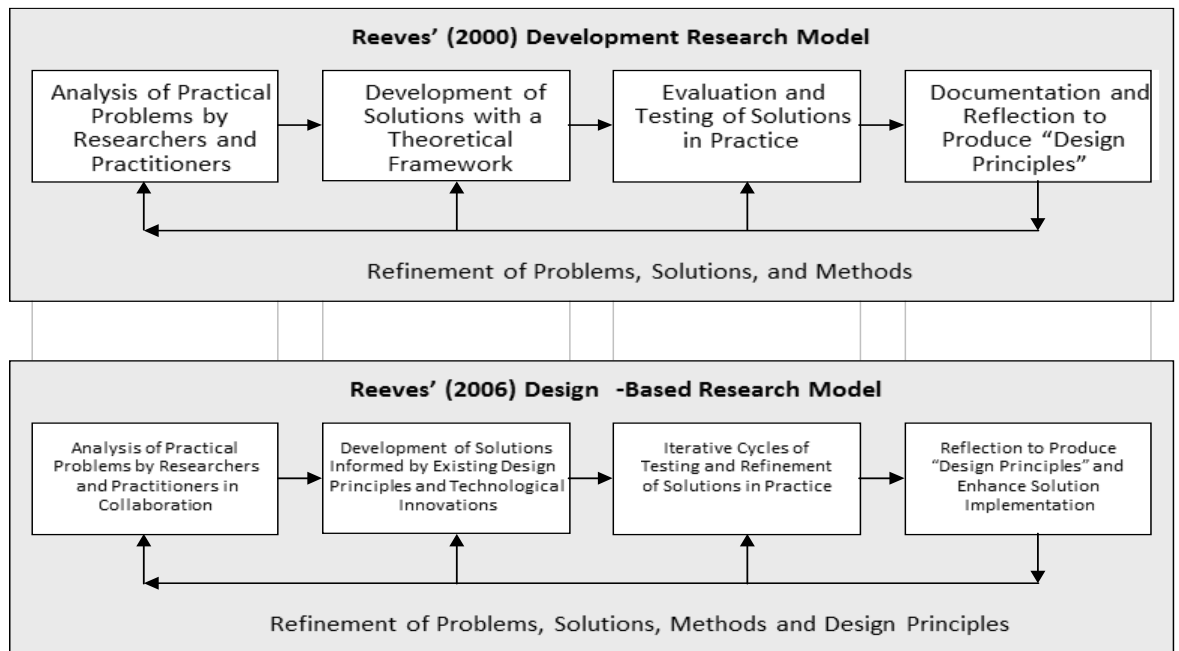


Figure (2) Reeves models

As it can be noticed from these models, design-based research is influenced by many fields including instructional design, evaluation and engineering. It is clear from Reeves models that design-based research is influenced by the famous instructional model “ADDIE (Analysis, Design, Development, Implementation, and Evaluation). However, “DBR moves beyond instructional design as craft knowledge towards understanding the know-how/know-why of the design” (Kelly, et al. 2008: 5). Sandoval & Bell (2004) refer to such influence indicating that “design-based researchers draw from multiple disciplines including developmental psychology, cognitive sciences, learning sciences, anthropology, and sociology. On the design side of the work, researchers draw from the fields of computer Science, curriculum theory, and instructional design and teacher education” (Sandoval & Bell 2004: 200). The main phases are usually either three, as in Richey & Klein (2005), which usually include product design and development, the product evaluation and the validation of designed tool, or four as presented by McKenney & Reeves (2012) model which include analysis & exploration, design & construction, evaluation & reflection and implementation & spread. Other researchers divide design-based research into different phases. For example, Easterday et. al. (2014: 3) specified six “iterative phases in which designers focus the problem, understand the problem, define goals, conceive the outline of a solution, build the solution, and test the solution”. Other models include Wademan’s (2007) model with five phases, McKenney’s (2001) model with three stages and Mafumiko’s (2006) model, also with three stages. Bannan-

Ritland (2003 cited in MacKellar 2010: 21) introduces a slightly richer, four-part model that has been drawn from other design fields and which entails the following four stages: 1) Informed Exploration; 2) Enactment; 3) Evaluation: Local Impact; and 4) Evaluation: Broader Impact. The same phases are discussed by Sahasrabudhe et.al. (2013: 3) and they are given different terms where the analysis phase is called “preliminary research: need and context analysis is done in the beginning,” the design phase the “prototyping phase: products/artifacts are created to address the problem and the evaluation phase as an “assessment phase: the intervention is evaluated to see if it addresses the problems and gives the desired outcomes”.

The model designed by McKenney and Reeves (2012) will be used in this study with its three main phases (analysis & exploration, design & construction and evaluation & reflection) besides an ‘optional’ fourth phase (implementation and spread) as illustrated in the figure (3) below. This model has been selected because the fourth phase (dissemination & spread) is noncompulsory which creates more flexibility for researchers who are interested in using DBR methodology. For some PhD candidates, this phase of the DBR processes cannot be achieved within the specified period or the allocated time for the research project as it entails procedures which take place after finishing the design of the instrument, the product or the programme. In other words, this model does not force the researchers to include summative evaluation, which may not be required or applicable to some studies due to time and funding constraints. Moreover, in some studies like this one, the last stage of formative evaluation or review usually ends with the required field tests, forming a kind of summative review that examines the product effectiveness and practicality to its potential users. Thus, in order to design and develop the checklist for the evaluation of teaching materials in the English Foundation Programmes, this model with its three core phases will be exploited. The first phase (analysis of the problem) usually begins from literature review till the full development of the study rationale and the second phase (development of solution) includes building the instrument based on the theoretical or conceptual frameworks that are shaped from phase one; the third phase (iterative cycles) comprises the implementation of the product or instrument as well as testing it with users and stakeholders. As it is clear from the model, the three phases lead to two outcomes: “theoretical understanding” of the whole project process and “maturing intervention” that is viable and beneficial and that can be used for the purposes specified by the researchers and stakeholders.

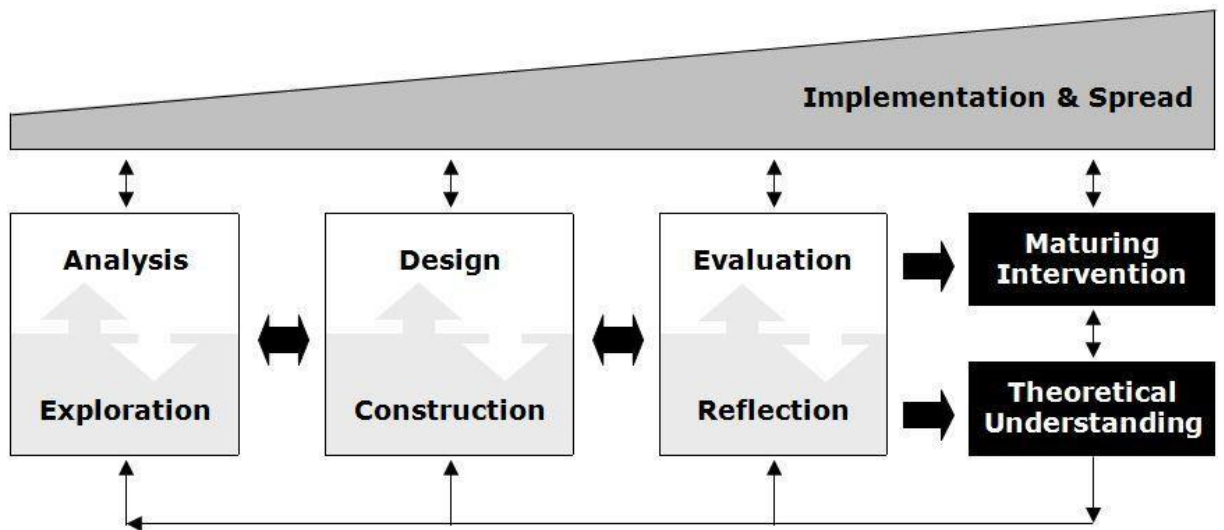


Figure (3) Generic Model for Conducting Design Research in Education. © 2012 McKenney and Reeves

This model is described in McKenney & Reeves (2012) as following:

The squares represent the three core phases, the arrows ...indicate that the process is iterative and flexible... the dual focus on theory and practice is made explicit through the rectangles which represent the scientific and practical outputs respectively...the trapezoid represents implementation and spread and the bidirectional arrows indicate that what happens in practice influences both the ongoing core processes and ultimate outputs (thus being contextually responsive) and vice versa.

In this study, the teaching materials evaluation checklist went through the three core phases illustrated in this model. These phases that are explained in the next section, comprise the (analysis and exploration phase) including literature review and context needs assessment, the construction of the checklist prototypes (second phase) and the systemic testing and the assessment of the developed checklist prototype by the potential users using four formative reviews and cycles (the third phase).

3.7 The Organization of the Three Core Phases in This Study

Based on (McKenney & Reeves 2012) model above, the first phase, consisted of three cycles using several instruments and activities which comprised literature review, informal discussions with six coordinators from Colleges of Applied Sciences, brainwriting sessions to collect data from students and teachers, a short survey, as well as experts' appraisal of the developed conceptual framework. As soon as all of main headings and sub-headings of the checklist were completed, and the conceptual

framework was validated phase two (design and construction phase) was initiated. So, the data collected in phase one along with the review of the literature resulted in specifying the main categories and the sub-categories of the checklist. In this phase (design and construction), four cycles were conducted, that included the development of the two initial prototypes, merging them into one major prototype, creating the checklist evolutionary prototype and developer screening of that prototype. The developed prototype of the evaluation checklist from phase two was then used in the (evaluation and reflection phase). This third phase consisted of four formative review cycles through the testing of the checklist with the different potential users (experts, teachers and coordinators) in the six Colleges of Applied Sciences besides experts from other higher education institutions in Oman (all the phases and cycles are illustrated in appendix F). In any design-based research study, four issues are important to facilitate its different phases and iterations. These are the research questions (discussed by Bakker, 2014), the general study plan (table 3 in chapter 1), the conceptual framework (figure 5 in chapter 4) and the DBR model (figure 3 in this chapter). Each one of these foci helps the researcher to concentrate on the different trajectories of this type of research, so that the whole processes lead to the findings that answer the research questions, yield a workable instrument and specify the instrument design guidelines and instructions of use. More details on these phases and cycles and on the procedures of the teaching materials evaluation checklist prototype development and assessment are demonstrated in the next chapters.

Chapter 4: The Development of the Teaching Materials Evaluation

Checklist

4.1 Introduction

This study was initiated as a result of several purposes represented by the five research questions in chapter 1. The core purpose of main question was to identify an appropriate method for designing a viable teaching materials evaluation instrument for the English Foundation Programmes in the Colleges of Applied Sciences in Oman. The conclusions from chapter two, demonstrated that evaluation checklists are the most appropriate and practical tools to evaluate teaching material in the English Foundation Programmes. Within the same chapter, it became clear that any evaluation checklist can emanate from two main sources: research and setting needs, which are the two primary foundations for the checklist developed in this study. A teaching materials evaluation checklist which is based only on one source, whether it is theoretical or contextual is considered inadequate for the English language programmes settings.

In this study, the researcher specified two main starting points for developing a practical and detailed teaching materials evaluation checklist. The first is the sources for the design of the checklist and the second is its main parts, categories and content. In other words, to make the development process easy and comprehensive, the designer has to start with a general concept about the design, which involves defining the main sources of the checklist based on either research or setting needs, or a combination of both, and after that decide about the headings and sub-headings of the checklist. The checklist developed in this study is based on both previous research & studies and practical needs. After deciding about the main categories of the teaching materials evaluation checklist, a conceptual framework was developed (see section 4.2.1). Then the conceptual framework was appraised by subject matter experts and as a result of the experts reviews, it was restructured based on their feedback. These processes are explained along with the main constructs of the developed checklist (research and practical needs) are a explicated in the following sections.

4.2 The Conceptual Framework

Clarifying the difference between theoretical frameworks and conceptual frameworks will help to understand the developing process of the study conceptual framework. Imenda (2014) proposes that a ‘theoretical framework’ refers to the theory/ theories

underpinning the research project or “pre-existing generalisation – such as Newton’s laws of motion, gas laws, that could be applied to a given research problem, deductively”. On the other hand, he defines the conceptual framework as the “end result of bringing together a number of related concepts to explain or predict a given event, or give a broader understanding of the phenomenon of interest-or simply of a research problem” (Imenda 2014: 189). It is the researcher’s responsibility to attempt to connect the different concepts of a design in a coherent procedure as “the researcher may have to “synthesize” the existing views in the literature concerning a given situation – both theoretical and from empirical findings” (Imenda2014: 189). So, developing a conceptual framework for the evaluation instrument design to link its different parts and categories was essential.

Previous frameworks for teaching materials evaluation checklists development were either too general or context specific. For example, Williams (1983: 1) discussed and proposed a framework on “how criteria can be developed for evaluating English language textbooks”, but at the end he could only present “a scheme for evaluation” and “instruction for using the checklist” with no explanation of how it is developed or validated. Another framework by Littlejohn proposes analysis at three levels: “what is there” in the teaching materials using “objective descriptions”, “what is required” by the “users” in the teaching materials, for example the content, tasks and learners’ roles, and finally “what is implied”, which requires “subjective inference” to discover the aims and principles of selection and sequencing (Littlejohn 2011, cited in McGrath 2013: 53). It has become clear that what is missing in the previous attempts by other researchers is a solid conceptualization of the whole development of the evaluation instruments. The early design of the framework, was basically originated from the various literature reviews conducted in phase one of this study (analysis and exploration phase) which resulted into specifications of the elements that should be considered when conducting teaching materials evaluations as detailed above. The initial prototype or design of the conceptual framework is illustrated in figure (4). It was later sent to four subject matter experts for validation and feedback.

4.2.1 Developing the Conceptual Framework

The framework is based on two main categories that comprise theory and practice. In the theoretical section, the categories that were specified earlier are: the Second Language Acquisition principles (what is the materials content to be evaluated); the teaching principles (how this content is delivered through textbooks) and the ELT curriculum

design principles (the practical means or ways that are used to combine content and teaching principles in the teaching materials). There are also three categories in the setting needs main category. These categories consisted of the potential users of the teaching materials in the English Foundation Programmes: the learners, the teachers and the educational institution. These categories are based on West (1994) who stated that there are three main parties where the needs are required to investigate. These are teachers, students and authorities. Thus, the needs of those three stakeholders were investigated through brainwriting data collection sessions and a short survey in order to complete the sources for the setting needs main categories of teaching materials evaluation checklist. Then they were incorporated into a conceptual framework along with the research main categories discussed above. The sub-categories of the two main sources or divisions (research and setting needs) will be explored thoroughly in the subsequent sections.

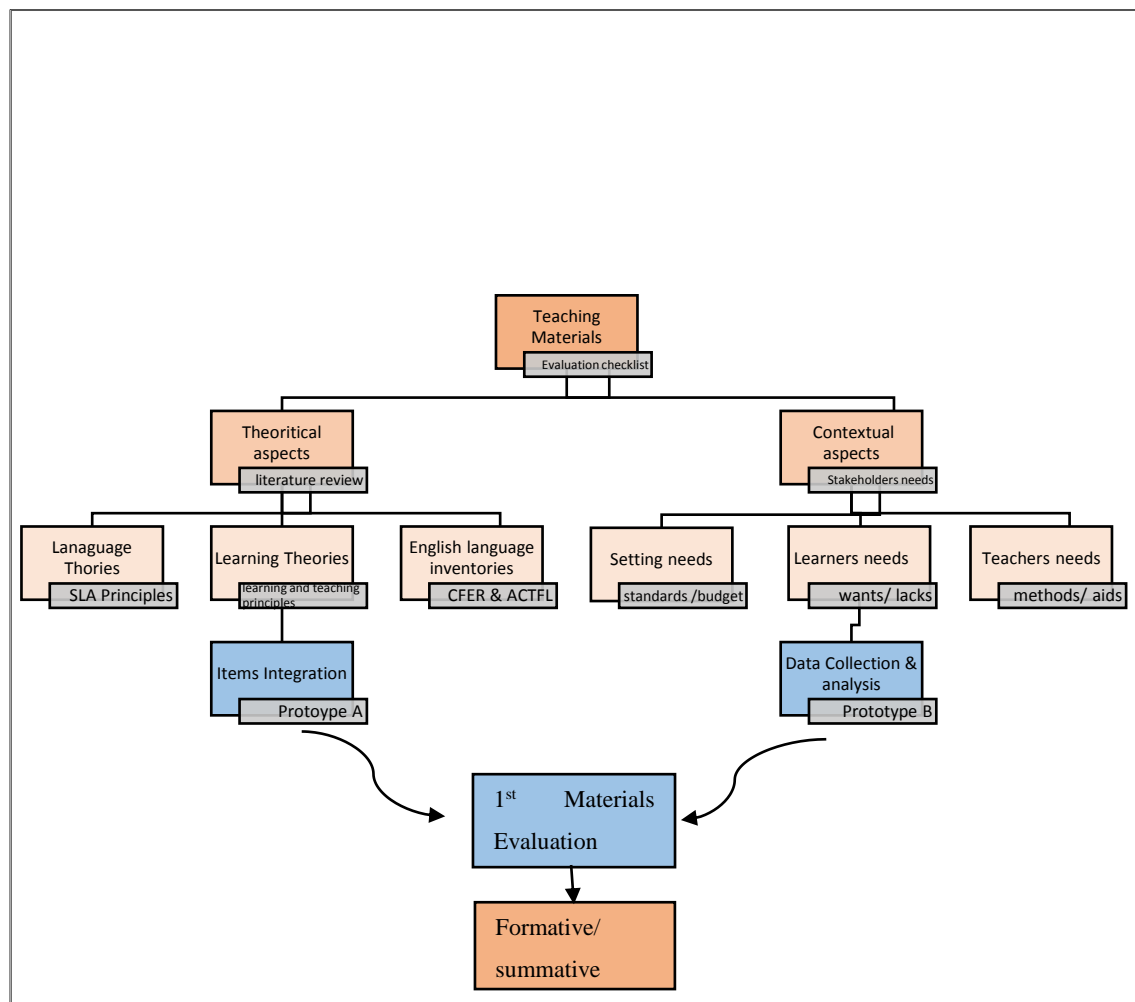


Figure (4) Conceptual Framework of the Sources for the Textbooks Evaluation Checklist in the English Language Programmes

4.2.2 Testing the Conceptual Framework by Subject Matter Experts

The developed framework went through many design processes in response to the feedback provided by four experts in the field of teaching materials development and evaluation. The goal of this expert review was to validate the conceptual framework before using it as the foundation and base of the detailed part of the teaching materials evaluation checklist design. To do so, “several evaluators are usually involved, as each individual typically only finds about one third of the problems” (Nielsen 1994 cited in Petrie & Bevan, 2009: 19). The following questions were sent to the four experts along with a short summary of the study and its main phases as well as the conceptual framework.

- 1) What is your first impression in terms of the framework’s practicality for target users (teachers, programme coordinators and experts)?
- 2) What do you think of the procedures of the framework development that led to the development of the 1st prototype of the teaching materials checklist?
- 3) What are the items that you think should be deleted or changed? Why?
- 4) What are the missing points or stages that you think should be included in the framework? Why do you think they are important?
- 5) What are the items or the processes in the framework that you think are not clear? What are your suggestions to improve them?

Based on the feedback received from the experts through answering the above questions, the framework was revised and modified to avoid the problems that may impede its understanding by the potential developers and users. Some excerpts of their valuable recommendations include:

There should be lines connecting the blue boxes to all three boxes above them to show that each of the prototypes is based on the three sources. The box titled “Setting needs” is not clear. (Expert 1)

The feedback from the experts was invaluable as it helped to make the conceptual framework more robust theoretically and practically. More detailed feedback was demonstrated through the comments provided by (Expert 2) below.

What is really missing is some indication of how all the information gathered can be combined when formulating evaluation criteria. How, for example, can the data on teachers’ needs be combined with information about learning theories? When the evaluator actually writes the checklist what categories are used and what is prioritized? Establishing a set of criteria for evaluating the evaluation instruments would be a very useful additional stage. How do the evaluators select,

from the mass of data, exactly what should inform their checklists? If everything is included, will the checklist be so unwieldy as to be unusable? I think you need to indicate how the framework should be used. Should the evaluators work through all the stages on one side first and then work down the other side or should they work across? If the latter, does it matter if it's from left to right or from right to left? Numbering the recommended sequence would help. I think you also need to indicate that using this framework isn't just a straightforward progression from stage to stage. It's a recursive process involving going both forwards and backwards and making numerous revisions.

The layout of the framework and the overall organization was tackled by (Expert 3) who suggested that it should be “circular” not “linear”:

When I first saw this framework, I asked myself, “why is it linear”? To my experience, instrument development is circular (allowing iterations), or to be more precise, heuristic and recursive. (Expert 3)

The next recommendations for improvement were from Expert 4, who focused on the selection of words and expression used by the researcher in the framework. He suggested clarifications of some “components”:

Perhaps Language Theories and Language Learning Theories can be combined otherwise you may be stretching this too far. Also, it is very unclear what is meant by “English Language Inventories” and what is the purpose of having it in the framework. I do notice that the justifications for the inclusion of all these components in the Framework are not explained very clearly. (Expert 4)

The expert appraisal of the conceptual framework was very informative, as the experts helped, through their expertise, to find out, at an early stage, the problems of the conceptual framework which will be used to design the closed or detailed teaching materials evaluation checklist. Such appraisal helped to “identify as many accessibility and usability issues as possible in order to eliminate them before conducting user-based evaluations” (Petrie & Bevan 2009: 18). The detailed feedback of the four experts on the design of the conceptual framework is presented in Appendix (C2). The conceptual framework was redesigned and some of the items were deleted, others added or amended as it is clear in the succeeding illustration of the revised framework (Figure 5).

4.2.3 Restructuring the Framework

Based on subject matter experts' valuable feedback, the conceptual framework was redesigned so that it can be easily understood and used in the development of the teaching materials evaluation checklist in this study as well as by other interested evaluation instruments developers.

Through this conceptual framework, the processes of the checklist development were made clear to the users and developers. Also, these procedures will help to simplify and facilitate the design of the teaching materials evaluation. The rearranged framework went from procedural to cyclical to indicate that the development of any product including the evaluation checklists should enable the designers to go through their design recursively. In so doing, developers can create more than one prototype and change it according to the feedback they get from the experts or the potential users. Some of the confusing terms were also changed to make the framework applicable to different users and stakeholders, especially teachers. The revised framework is more detailed comprehensive and easy to follow. All the terms used are simple and clear for almost all English Foundation Programmes practitioners. The processes of the teaching materials evaluation checklist development can also be traced and followed easily through the amended conceptual framework. Within this framework, the evaluation checklist developer is left with three options for the instruments sources. The developer can design a checklist that is based on theoretical sources only, on the local or contextual sources or a combination of both sources and components. The new revised conceptual framework is illustrated in (Figure 5) below.

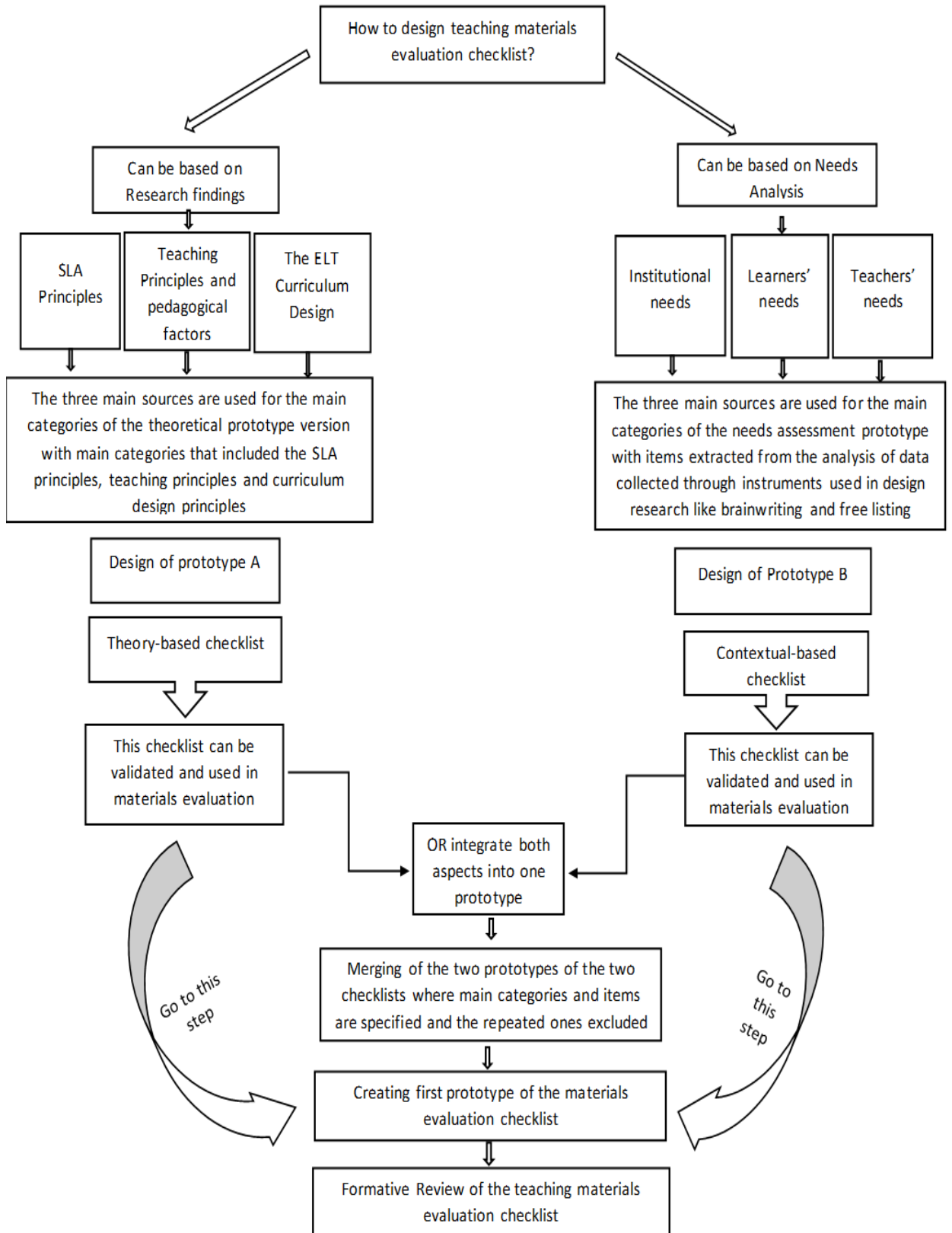


Figure (5) The Revised Conceptual Framework of the Sources for Textbooks Evaluation Checklist in the English Language Programmes

4.3 The Main Constructs of the Checklist: Research

Research on the development of teaching materials evaluation and particularly checklists (in previous studies) utilizes different sources, and each evaluator usually selects one of the sources as the main base of the checklist development. The tradition where the “the evaluators’ theory of learning and teaching” (Tomlinson 2003: 17) is considered the inception of checklists development without any generic framework or model to guide the design processes, has led to a perplexing situation. The result is myriad of checklists which are mostly applicable only by their developers with no reciprocal conceptions or geneses. Therefore, defining specific and unified sources from research will contribute towards a principled approach for the design of the teaching materials evaluation checklists. As it is obvious above in the conceptual framework that will be used in the checklist development, each main source (research and setting needs) is divided into three main categories: SLA principles, teaching principles and ELT curriculum design principles in the research aspect, and students’ needs, teachers’ needs and institutional needs in the setting needs aspect. Each of the sub-categories has several items that describe the features and qualities that the evaluators will have to check when developing a checklist for choosing or evaluating teaching materials in the English Foundation Programmes. The three main categories that were based one extensive review of literature and that were thought to be the appropriate basis for teaching materials evaluations checklists will be discussed next.

4.3.1 Second Language Acquisition Principles

Language acquisition, whether first or second language, has been the subject for numerous studies, theories and approaches which try to explain how children and learners acquire languages. So, as there are many learning theories and models in the literature, Second Language Acquisition principles were selected to be part of the research aspect of the teaching materials evaluation checklist as SLA can be considered the summary of most if not all of these educational theories. Saville-Troike (2012: 2) defined Second Language Acquisition (SLA) as “the study of individuals and groups who are learning a language subsequent to learning their first one as young children, and to the process of learning that language” which indicate the importance of SLA for materials development as “L2 materials developers can make use of SLA research to help them to develop principled materials which can facilitate the acquisition of an L2 in the classroom” (Tomlinson 2013 b: 25). As these principles become part of materials development, their

role in the materials evaluation also becomes essential. Menezes (2013), thinks that despite the “huge number of SLA theories and hypotheses” mostly “eight of them: behaviorism, acculturation, universal grammar hypothesis, comprehension hypothesis, interaction hypothesis, output hypothesis, sociocultural theory and connectionism” have “caused more impact” in education (Menezes 2013: 404). These theories were usually distilled into general principles or models for use and practical application in the learning and teaching contexts. According to Myles (2011), two main findings have affected second language learning: “second language acquisition is highly systematic” and “second language acquisition is highly variable”. For the first point, many research models have been developed to explain the developmental processes to acquire second languages, such as “Universal Grammar, Cognitive models, Interactionist/ Sociocultural models” (Myles 2011: 5). These models propose that language acquisition results from extensive input and social interactions, implying certain principles and methods for teaching, which will be explained later. Regarding the variability aspect, and despite the findings in SLA which recommend that learning is a systemic process, variables “in route, rate and outcome” of learning will affect the learners’ success. Thus, whether the variables are internal, such as the effect of first language as in language “transfer”, or external, such as “intelligence, aptitude, motivation, attitude... and social and sociolinguistic” (Myles 2011:11), the differences between learners in different contexts should be considered when developing and evaluating English language teaching materials. Second language acquisition research has influenced the field of English language learning and teaching, proposing certain techniques for both learners and teachers. Rubdy (2003) explains how Tomlinson (1998) recognized the importance of Second Language Acquisition use in materials development and evaluation. Rubdy pointed out to that through referring to Tomlinson’s introduction about materials development indicating that “many of the tenets and basic principles of second language acquisition” can be used as indicators to “understanding of what good materials” are besides using them as “principles judgment about” the teaching materials (Rubdy 2003: 43). Therefore, the principles of SLA can be of great importance when incorporated in the development of the evaluation instruments.

Saville-Troike (2012: 5) examined Second Language Acquisition from “three different disciplinary perspectives: linguistic, psychological, and social”, trying to answer three questions: “what exactly does the L2 learner come to know? How does the learner acquire

this knowledge? Why are some learners more (or less) successful than others?” Answering such questions through SLA research and findings that made use of different disciplines (linguistic, psychology and sociology) can help to design effective materials as well as effective evaluation instruments as such information can be used as a guide for developers and evaluators. Thus, despite criticism, general principles derived from these theories can form a solid ground to develop teaching materials, design different tasks and evaluate teaching materials and the quality of activities and exercises. Books and articles have been published studying the influence and findings of SLA on materials development and evaluation (Harwood 2010, Tomlinson 2011, 2012, 2013, a & b, McDonough et. al. 2013). Despite that, understanding its principles is not usually communicated well to practitioners, so incorporating them in the evaluation checklist will facilitate their use and understanding.

Developing a viable checklist for teaching materials evaluation, should consider such language learning and acquisition principles to make sure that the textbooks are based on well-established findings of research, and not only on the authors’ single experience of teaching or learning foreign languages. Thus, fifteen SLA principles were selected, in the first instance, to be included in the teaching materials evaluation checklist initial prototype. As a start, the items selected for this sub-category of the research aspect in the checklist were based on the work and analysis of experienced researchers in the development of evaluation checklists for teaching materials, such as McGrath (2002), Tomlinson & Masuhara (2014), Tomlinson (2013) Adaskou et.al. (1990) and Ur (1990). Throughout the different phases and cycles of the checklist review by different experts and stakeholders, the six most desirable principles were kept in the final checklist prototype. It can be seen that the six principles are borrowed from different SLA theory sources. The first four Principles are based on social theories such as “Acculturation theory and social Psychology...and sociocultural theory” that focus on how “factors as identity, status, and values affect the outcomes” of Second Language Acquisition (Saville-Troike 2012: 29). The six principles that were based on Tomlinson’s and the mentioned researchers above are:

- 1) Materials should help learners to gain confidence (Dulay, Burt and Krashen 1982 cited in Tomlinson 1998)
- 2) Materials should provide the learners with opportunities to use the target language to achieve communication purposes (Swain

- 1985, 2005)... to develop strategic competence (Canale and Swain, 1980 cited in Tomlinson 2013)
- 3) Materials should take into account that the learners differ in affective attitudes. (Gardner and Lambert, 1972; Naiman et al, 1978; Crookes and Schmidt, 1991; Gardner and MacIntyre, 1993, Oxford and Shearin, 1994; Dörnyei, 2001 cited in Hurd, S. (2002).
 - 4) Materials should help the learner to develop cultural awareness (Byram and Fleming, 1998; Tomlinson 2000 cited in Tomlinson 2014)
 - 5) Materials should cater for the needs of diverse learners (Tomlinson 1998)
 - 6) Materials should help learners to personalize their learning (Tomlinson, B., & Masuhara, H. (2013)

The above principles formed the source for the items in the second language acquisition principles sub-category. Each principle was simplified for the evaluator through clear explanations and instructions on what exactly they should search for when evaluating teaching materials. This characteristic was intended to make the checklist's understanding and use stress-free especially for evaluators as inexperienced teachers and programme coordinators. On the empirical aspect of SLA principles, Tomlinson (2013) makes use of these principles to design his own criteria for the development and evaluation of teaching materials. The principles used are:

“a rich and meaningful exposure to language in use, Affective and cognitive engagement, making use of those mental resources typically used in communication in the L1, Noticing how the L2 is used, Being given opportunities for contextualized and purposeful communication in the L2, being encourage to interact, being allowed to focus on meaning”
Tomlinson, (2013: 99)

Most of the above principles are derived from theories of language acquisitions that had been studied as stated in Ur (2012: 6) such as Intuitive acquisition where “we learn another language the same way as we learnt our first...through lots of exposure to the language in authentic communicative situations (Krashen, 1982); Cognitive process where “language involves the understanding of underlying rules: if we master these rules, we will be able to apply them in different contexts (based on Chomsky, 1957) and Skill learning where “language is considered a skill so we learn in in school just as we learn other skills: someone explains rules or words to us, we understand and practice them until we master them and use them fluently and skillfully” (Johnson, 1996). Masuhara, H., & Tomlinson (2008: 23- 36), is an example of SLA practical studies that involved using

SLA principles in the evaluation of English language teaching materials. They used principles of SLA to evaluate seven coursebooks for general English used in UK. Through their evaluation, they were able to specify the strengths and the weaknesses of the teaching materials. Both evaluators also provided suggestions for the improvements of these materials. The next sub-category related to research-based aspect is teaching principles or the pedagogical approaches that are usually considered in the procedures of teaching materials development, use and evaluation.

4.3.2 Teaching Principles

In the field of English language teaching, the search for best practice has led to constant changes of methods and approaches of teaching. Teaching principles are referred to by some researchers as teaching methods or approaches (Brown 1980: 240) and teaching theories (Knowles 1973: 66), pedagogical factors or teaching styles. Nowadays, teaching materials that are based on certain types of syllabuses (usually communicative or mixed syllabus) still follow certain teaching models like the famous PPP “presentation-practice-production”, the ESA that “refers to the engage-study-activate” (Harmer, 1998) or the PPT “presentation, practice and testing” model (Ur 1996) as elucidated by Zhu & Liao (2008: 92). This format and presentation of textbook organization is supposed to go through major changes as a result of the various developments and variations of learning and teaching conceptions. The history of teaching started with pre-methods, where simple approaches of teaching were used and every teacher used his or her own way of instruction. Then, the era of methods emerged as a means to organize the practice of teaching. Several methods were introduced, based on different theories of learning and teaching the language. ‘Method’ here is identified as an “overall plan for presentation of teaching materials based on a selected approach” (Tasnimi 2014: 2). This period was “between the late 19th century and the late 20th century” (Maghsoudi 2016: 283). Many teaching methods were advocated, from the Grammar Translation Method, the Direct Method, the Audiolingual Method, the Silent Way, Suggestopedia, Total-Physical Response, Community Language Learning, and Communicative Language Teaching (CLT) to Content-Based Instruction and Task-Based Language learning as alternatives to CLT. All of these methods have been criticized as there was no single method that can be identified as the most appropriate one for learning and teaching second languages. In consequence, and “after the successive rise and fall of a series of methods and approaches in the early and mid-twentieth century, the English Language Teaching (ELT) researchers

and practitioners came to realize that no single method or approach of language teaching would be the optimal framework to guarantee success in teaching a foreign language” (Fat’hi et. al. 2015: 306). Pennycook (1989) argued that methods actually serve the dominant power structures in society, leading to “a de-skilling of the role of teachers, and greater institutional control over classroom practice” (Pennycook 1989: 610, cited in Islam, 2017: 540). As a result, eclecticism emerged within Communicative Language Teaching “as a desirable, coherent, and pluralistic approach to CLT” and where “principled eclecticism involved using different language attitudes that have different characteristics in response to learners’ needs” (Maghsoudi 2016: 283-284). Soon this concept about teaching English was rejected as being theoretically unproven. Finally the occurrence of post-method era resulted in the “death of methods” (Allwright 2003). Following this, “reflective teaching was introduced because it promised to be the solution to the dilemma” (Akbari, 2007, cited in Fat’hi et. al. 2015: 307) where “practitioners are allowed to theorize from their practice and practice what they have theorized” (Kumaravadivelu, 1994: 30). Therefore, “they are not supposed to follow in the footsteps of any teaching gurus” (Fat’hi et. al. 2015: 312). Knowles has led another movement, coining the term ‘andragogy’ as opposed to pedagogy where “Knowles’ concept of andragogy - ‘the art and science of helping adults learn’ - ‘is built upon two central, defining attributes: first, a conception of learners as self-directed and autonomous; and second, a conception of the role of the teacher as facilitator of learning rather than presenter of content” (Pratt & Ass 1998: 12) and this emphasizes “learner choice more than expert control” (Reischmann 2004: 3). Knowles’ concepts about the teacher as facilitator requires more attention to teachers training programmes. This approach can be helpful in English language programmes in tertiary education as the learners in these institutions will need special teaching techniques that suit their age and their interests. Most of the learners in the English Foundation Programmes are at the age of eighteen and nineteen. For that reason, teaching approaches used in the teaching materials development and evaluation in English Foundation Programmes are supposed to consider these changes as well as learners’ age and teachers’ professional development.

Teaching principles and pedagogical factors selected for the teaching materials evaluation checklist in this study, were first based on Nation & Macalister (2010) and Ellis (2005) attempting to use general statements about the techniques of imparting knowledge to the learners rather than following certain methods or approaches that had been already

abandoned. Those scholars' views and interpretations were chosen because they tried to separate teaching principles from teaching methods. Later, and through the various revisions of the checklist prototype, only three teaching principles were kept in the final version. The three principles in this section of the evaluation checklist were based on research that exhibited how can teachers help students learn to connect previous information to new information or "making use of the "students' built in syllabus" as well as "extensive use of L2 and extensive L2 input" (Ellis, 2005). So teaching principles in this section of the teaching materials checklist is basically focused on empowering teachers through certain conceptions about teaching, where they can move from following the textbooks literally, in the "science-research conceptions" and "theory-philosophy conceptions", to the innovative teaching principles of "art-craft conceptions" in which the teacher is able to envision the materials he/she is teaching and is capable of mining the important information in the textbooks and supplementing for the missing items and language content needed by their learners. With such developments, teachers' roles have changed into influential facilitators and participants in the educational policies and plans. In fact, the "post-method teachers are autonomous, analysts, strategic researchers and decision-makers. Such teachers are also reflective as they observe their teaching, evaluate the results, identify problems, find solutions, and try new techniques" (Fat'hi et. al. 2015: 309). Based on all of these changes, there is a movement from "science-research conceptions" towards "art-craft conception of teaching" (Arikan, 2006: 4) as well as a shift from top-down process to bottom-up process to allow teachers more freedom to improve their instruction and their experience.

To explain the conceptions about teaching theories mentioned above, Zahorik (1986) divides teachers into three categories according to their daily practice of good teaching. His concepts of good teaching included "Science Research Conceptions, Theory - Philosophy Conceptions and Art-Craft Conceptions" (Zahorik 1986: 22-23). According to the principles of the first approach, "science research conceptions", teachers depend totally on other resources in their instruction. These are derived from three sources: "doing what effective teachers do, following a tested model and operationalizing learning principles". The second approach to sources, "theory-philosophy conceptions" is based on "implementing theoretical models" and "a philosophical model". So, again the teachers here depend on the provided models of teaching in the textbooks. Finally in the "art-craft conception", the teacher can perform "in resourceful, creative ways". Richards

(2002: 19-25) explains Zahorik's conceptions about what is considered to be "good" teaching principles and their origins. He then summarizes the principles that teachers should follow according to the three conceptions to include principles like: monitoring "students' performance on tasks to see that desired performance is being achieved" in science-research conceptions, observing teacher's "teaching to see that it conforms to the theory", in theory-philosophy conceptions and developing "personal approaches to teaching" in the art-craft conceptions. As it becomes clear, English language teaching materials usually 'deskill' teachers through keeping them uninformed about the principles of good teaching, which may impede communicating knowledge to their learners and 'reskilling' themselves at the same time.

Zahorik's conceptions of teaching raise questions about teachers' guides or manuals and if they ever consider teacher education programmes and their professional development when they are designed. In most cases if not all, teachers' guides can be considered answer keys to the materials' activities and exercises with no further intentions to encourage innovative techniques in delivering the materials' content. These manuals are supposed to reflect the teachers' needs in their development process from a "developing teacher" or novice teacher (represented by Zahorik's first two conceptions) to "experienced teacher" (represented by Zahorik's third conception). Later, in this study, and through the data analysis, these types of teachers will be noted, who differ in their evaluation of the teaching materials using the developed checklist, where their comments on the checklist items may look contradictory (e.g. an expert asks for the use of simple terms and an inexperienced teacher will demand the use of theoretical terms as he is familiar with such terms in his teacher education programme). From the above discussion about teaching conceptions, only the substantial principles in the published teaching materials can be evaluated. Thus, the three principles used are expected to help teachers not only to evaluate their teaching materials, but also to teach skillfully, to understand what they are doing and to increase their expertise.

The principles remained in the checklist revolve around assisting the teachers to help learners to make the most effective use of the previous knowledge, "science-research conceptions", the availability of enough sources and guidance to provide extensive use of second language in and outside the classroom, "theory-philosophy conceptions" and finally, the instances offered to teachers to provide opportunities for learners' language productions, also in "science-research conceptions". Some of the items in SLA and

teaching principles appear to be similar due to the multiple uses of second language acquisition theories and findings which led to some indistinct boundaries between the three main areas used in teaching materials development and evaluation and used in this study that include: learning principles, teaching principles and ELT principles. This may justify the repetition of some items in any of the three areas. In fact, the studies using second language acquisition “has informed the work of syllabus designers, methodologists, and materials writers” Nunan (2001: 91).

Despite similarities, the section of the teaching materials evaluation checklist on teaching principles is different from teachers’ needs, which is discussed later in setting needs. The teaching principles here are meant to facilitate instructors’ teaching through helping them to provide the most useful resources (from research) to ensure the continuity of students’ progression and achievements. The selected principles are related to science-research conceptions and theory-philosophy conceptions, as most teachers’ experiences in the language programmes range between those two concepts and because these are the dominant features in the commercial textbooks. Also the majority of commercial textbooks writers’ focus is to help teachers in those two categories in their daily practice of teaching in English Foundation Programmes. The art-craft conceptions are not acknowledged in these language programmes nor in the teaching materials, as teachers are not given the freedom to use their own materials, their own approaches or their own types of assessments. Despite that, it is hoped that textbooks writers will seriously consider the art-craft conceptions, which will give more options to teachers and encourage them to be creative in their instruction. The development and the promotion of the teaching material evaluation instruments design and use are also techniques for raising teachers’ awareness as well as their professional development.

As is noticeable in the ELT field, the disappearance of methods, led to the abandonment of single syllabuses which have usually been replaced by mixed syllabuses and general frameworks for materials development and evaluation. Hence ‘the ELT curriculum’ is the third heading in the research-based source of the teaching materials evaluation checklist.

4.3.3 The ELT Curriculum Design

The move from the 'methods' period to the post-method era initiated changes in the design of the ELT curriculum besides the method of instruction. Thus, the ELT Curriculum

design came across several modifications departing from “the information-oriented system into an inquiry-oriented” one (Tasnimi 2014: 7). Earlier models of curriculum such as Tyler’s model are basically based on behavioural objectives, where he recommends methods for the specifying the “educational purposes”, the “learning experiences” to obtain these purposes, and the organization of “learning experiences...for effective instruction” as well as suggesting a method for evaluating “the effectiveness of learning experiences.” (Tyler 1949: 104). Stenhouse’s model is based on seeing curriculum as a process rather than a product and thus he considers curriculum as “an attempt to communicate the essential principles and features of an educational proposal in such a form that it is open to critical scrutiny and capable of effective translation into practice” (Stenhouse 1975: 4). In these models, evaluating a curriculum is equivalent to evaluating a textbook. Which may mean that the evaluation of materials should consider experiences and skills as “study skills, arithmetic skills, and social skills” and values like “the dignity and worth of every individual regardless of his race, religion, occupation, nationality, or social class” (Bellack & Kliebar 1977: 51). According to this definition, curriculum is stretched to comprise extracurricular activities, cultures, learning environment, and a hidden curriculum alongside the main components which include content, goals, methods and assessment. Richards (2013) tries to narrow the elements of the ELT curriculum by identifying three main types of curriculum approaches: “forward design” that “starts with syllabus planning, moves to methodology, and is followed by assessment of learning outcomes”; “central design” that “begins with classroom processes and methodology” and where “syllabus and learning outcomes” are left to be “addressed as the curriculum is implemented”. In the central design the teaching materials developers do not have to “clearly defined outcomes or objectives” because “learning takes place in a context and evolves through the interaction and participation of the participants in that context” so the syllabus design becomes cyclic where each process feeds the other and based on that the objectives, the content and the assessment can be identified .Finally the “backward design” that “starts from a specification of learning outcomes” where “decisions on methodology and syllabus are developed from the learning outcomes” (Richards, 2013: 8-28) . The processes of curriculum design were also connected with new technological approaches and to instructional design. Acun (2011: 834), mentions a third approach in comparison to the process and product approaches:

There are two main approaches to developing a curriculum: the product approach proposed by Ralph Tyler (1949) and the process approach usually associated with Lawrence Stenhouse (1975). The systems approach, which originates from the computer systems, is emerging as a third main approach due to, perhaps, the spread of computer systems in all facets of life.

In the systems approach, or instructional design, the focus is on teaching principles more than on learning principles. Many instructional models were developed for materials design: “several models are suitable for the design of instruction of course units and lessons. One widely known model is the Dick and Carey” (Gagne & Briggs 1974: 21). Gagné & Briggs and Dick & Carey models are the most famous and mostly consist of five stages: analysis, design, development, implementation and evaluation. In fact, “instructional designers” as a term is considered the substitute for curriculum designers. But as they are “unable to completely identify with Tyler’s rationale, instructional designers contrived an ID rationale” (Kemp et.al.1996: 4, cited in Petrina 2004: 90-91) where they have to think through these general questions:

- 1) For whom is the program developed? (characteristics of learners or trainees)
- 2) What do you want the learners or trainees to learn or demonstrate? (objectives)
- 3) How the subject or skill is best learned? (instructional strategies)
- 4) How do you determine the extent to which learning is achieved? (Evaluation Procedures).

Besides the technological advances, “the three shifts—from communicative language teaching to task-based language teaching, from method-based pedagogy to post-method pedagogy, and from systemic discovery to critical discourse—constitute the major transition in TESOL methods during the past 15 years” (Kumaravadivelu 2006 a: 71) and have affected the design and the evaluation of the curriculum as well as teaching materials. Thus, the change of methods and approaches of teaching has led to further changes in the ELT curriculum design. Kumaravadivelu highlights “three postmethod frameworks: (a) Stern’s three-dimensional framework, (b) Allwright’s Exploratory Practice framework, and (c) Kumaravadivelu’s macrostrategic framework” (Kumaravadivelu 2006 b: 185). Post method pedagogy in Kumaravadivelu's framework defines three guiding principles. These are Particularity, Practicality and Possibility where Particularity pays more “attention to local contingencies” (Fat’hi et. al. 2015: 312) rather than depending on the external experts and policies. Practicality “focuses on

teachers' reflection and action, which are also based on their insights and intuition" as teachers are able to use their "prior and ongoing experience with learning and teaching" to "gather an unexplained and sometimes explainable awareness of what constitutes good teaching" (Kumaravadivelu 2006 b: 173). Finally Possibility where "L2 teaching is seen more as a tool to help learners come to grips with their own identity and as a vehicle to explore other peoples and cultures" so second language learners within this principle will attempt to acquire "not just a new linguistic experience but, more importantly, a new lens through which to appreciate the world out there and the world inside, hence the global and local becoming part and parcel of the whole L2 experience" (Fat'hi et. al. 2015: 312). Those assumptions were part of the critical pedagogy movement. Yet again, the "the criticism about research" conducted "in critical pedagogy could... be extended to research in TESOL in general and TESOL methods in particular" necessitating "the search for robust research design" (Kumaravadivelu 2006 a: 74-75) of different educational matters and concerns including teaching materials development and evaluation.

The shift from curriculum foundations mentioned above to general principles about the teaching materials requires establishing general criteria that can help in evaluating their effectiveness as well as their adaptation. But the question that remains unanswered is which point of view in relation to the ELT curriculum and teaching materials should be evaluated in the commercial textbooks and whether or not there is a general concept that can form a base for the evaluation of these materials. The above mentioned frameworks by Kumaravadivelu (2006b) are connected to certain post methods ideas, for example critical pedagogy. Therefore, in this study preference was given to more general approaches to ELT curriculum, and which have been accompanying the changes in learning and teaching English steadily with more caution and less radical transformations.

According to Brown (1995), three authors have contributed to the demarcation of the ELT curriculum: Anthony (1963), Richards & Rodgers (1982) and McKay (1978). Anthony's framework included "three categories: approach, method and technique" where the first refers to all "the points of view on the nature of language, and the nature of language teaching and learning" (Brown 1995: 140). Richards & Rodgers (1982) made use of Anthony's framework and kept the approach as the basis of the ELT curriculum theory, replacing the term 'method' with 'design' and 'technique' with 'procedures'. For McKay (1978), Brown explains how she divided the language syllabuses into also three main

types: “Structural syllabuses, situational syllabuses, and notional syllabuses”. In developing the checklist for the teaching materials evaluation in English Foundation Programmes, the ELT curriculum, pedagogical approaches and instructional principles have to be considered in order to create a viable and reliable checklist as McGrath (2002: 27 citing Tucker 1978:219) proposed, to base the criteria on “basic linguistic, psychological, and pedagogical principles underlying modern methods of language learning.” The items for the checklist in the ELT curriculum design are based on Richards & Rodgers (2014) framework of ELT curriculum, where they think that certain principles should be made noticeable in materials development and eventually in their evaluation. As a result, the ELT curriculum principles used in the development of the teaching materials evaluation checklist in this study include: (1) “the method or pedagogical approach of the materials” underlying their design and instruction should be “made clear to the users”, (2) “the use of pedagogical activities is well explained” to the users and (3) the “procedures and techniques in giving the feedback on the activities to the learners” should be also “explained” to the teachers (Richards & Rodgers 2014: 22-40). The focus on tasks in the teaching materials evaluation checklist was because of importance of task and activities in the second language learning and teaching as Prahbu (1987 explain “units of syllabus organization should be tasks which define which outcomes can be achieved through language rather than linguistic items, learning will be effective only if it is related closely to language use and involves relating form and meaning.” As a result of their importance, tasks clarification to the teachers becomes as important as tasks design.

From reviewing the research in the related literature, it is evident that the development of the teaching materials evaluation checklist is “no easy matter” (McGrath 2002: 43). At the same time it cannot be based on a single aspect of educational research or theories as this will affect its validity and reliability. It is rather a combination of research strands and the settings needs which should be addressed in the evaluation criteria. It is hoped that approaching the ELT curriculum through such principles will facilitate its development and evaluation as well as solve the dilemma of the theory/practice issue in the English Foundation Programmes. With the theoretical aspect determined in the teaching materials evaluation checklist, the setting needs will be specified next for the main categories and sub-categories of the checklist.

4.4 Main Constructs of the Checklist: Practical Needs

4.4.1 Teachers, Students and Institutional Needs

According to Johnston & Peterson (1994: 63), the stakeholders of English language programmes are: “learners, teachers, administrators and controlling authorities”. West (1994) in his answer to the question “who should decide what the language needs are?” states that there are “three principal parties involved in what has come to be called the needs analysis triangle” (West 1994: 6); these are the learners, the teachers and the educational institutions as illustrated in Figure 6. In any language programme, the needs of the stakeholders are essential for the success of the programme. Therefore, in English language programmes, the most important stakeholders are the users of the teaching materials: the learners, the teachers and the authorities or the institutions.

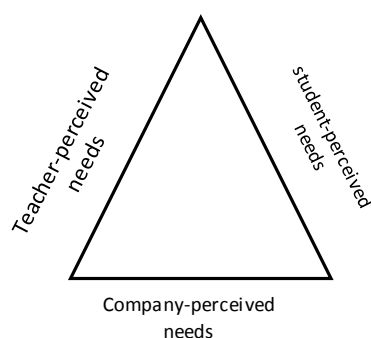


Figure (6) The needs analysis triangle based on West (1994)

Brown (1995: 21) defines needs analysis as “the systematic collection and analysis of all relevant information necessary to satisfy the language learning requirements of the students within the context of the particular institutions involved in the learning situation.” In response to this definition of needs, collecting data from the involved stakeholders in these programmes becomes an essential part of any teaching materials evaluation tool. Chostelidou (2011: 403) states that “the process of needs analysis has been established as a key concept of ESP course design, program implementation and evaluation” mentioning many scholars who investigated this issue, including Brown (1995), Dudley-Evans & St. John (1998), Ellis & Johnson (1994), Johns & Price-Machado (2001), Jordan (1997), Munby (1978), Richards (2001), Flowerdew & Peacock (2001), Hamp-Lyons (2001) and Boshier & Smalkoski (2002). In addition, “needs analysis tends to be associated with ESP, and is neglected in the General English classroom” (Seedhouse 1995: 59). In fact, “what distinguishes ESP from General English is not the existence of a need as such but rather an awareness of the need” (Hutchinson

and Waters 1987:53-4 cited in Seedhouse 1995: 59). Needs analysis studies are usually based on general feedback obtained by teaching materials developers through secondary sources. Also, materials are developed by international publishers who may have only a list of the sensitive or taboo topics besides their academic experiences in the field. These mass-produced teaching materials that come in packages do not provide any evaluation instruments or guidelines that may help the users to select the appropriate materials or to evaluate their effectiveness while or after use. Seedhouse (1995: 59), states that “for the time being, the tradition persists in General English that learners’ needs can’t be specified and as a result no attempt is usually made to discover learners’ true needs.” This assumption is incorrect, because the stakeholders' needs can be identified through many data collection methods that involve the use of innovative instruments and several iterations of data collection. One of the instruments used in this study is the ‘brainwriting’ technique which was used to explore the needs of teachers and students in the English Language Foundation Programme.

The teachers can be described as the delivery tool, the transmitters of information and knowledge and the managers and supervisors of learning using teaching materials as the medium (Gray: 2013). McGrath (2013) talks about the teachers’ roles regarding teaching materials and includes their contribution in choosing the materials, controlling the use of materials, and being creative in teaching these materials in the classrooms. They may use different strategies mentioned by Maley (2011) such as: omission, addition, reduction, and reordering of content if they have the opportunity to do so. It is important to recognize that the teachers’ needs here are related to their views about the content of the textbooks and their recommendations about the appropriate linguistic content and skills, rather than identifying their own individual needs and interests.

To discover the teachers’ views about teaching materials in the Colleges of Applied Sciences, brainwriting sessions were conducted with six teachers, which helped to identify their suggestions about the good features required in textbooks and the different elements recommended within the four main skills, as well as other aspects that can be investigated for materials selection and evaluation. The items resulting from these sessions (detailed in subsequent sections) were used in the headings and sub-headings in the teaching material evaluation checklist representing teachers’ opinions on these issues.

Students are the main users of the textbooks beside teachers. These textbooks are often their only source for learning English language in the classroom and at home, especially where English is considered a foreign language for most if not all of them, and where the chances to practise outside the educational institution are not available. In selecting and evaluating teaching materials in the English Foundation Programmes, issues such as the purposes of learning the language, the effects of the environment on their learning styles and learning strategies should be considered, and this is what was explored through the brainwring sessions with 24 students in the Foundation Programme. The results of these sessions formed the categories and items of students' in the development of the checklist for the teaching materials selection and evaluation for the English Foundation Programme. Again, it is important to know that the needs here are about the qualities of the textbooks, not the subjective needs of the learners, which can be studied in contexts other than this study.

In choosing teaching materials, decision makers' main concerns, beside the quality of materials, are financial and availability issues, in addition to the alignment of these materials to the institution's mission, visions and strategic objectives and the standards of the educational institutions. In English Foundation Programmes which are designed to enable students to reach certain proficiency levels to study specific subjects and to master certain skills to become successful local-global citizens, teaching materials are supposed to be selected and chosen to align with such goals and aims. In the Colleges of Applied Sciences, there are specific National Standards which all English Foundation Programmes are expected to consider when selecting any teaching materials, as discussed in chapter 1. The standards are basically for (A) level students, who are supposed to be on the appropriate proficiency level before they can proceed to their degree programmes. To specify some of the important considerations, for the authorities in the Colleges of Applied Sciences, a short survey was sent to one of the policy experts in the Ministry of Higher Education. The results of that survey were incorporated in the developed checklist under the category of institutional needs. The setting needs assessment procedures are explored through the succeeding sections.

4.4.2 The Instruments Used for Needs Assessment

The first phase of the design research involved the analysis and exploration of the context. This phase included several formal and informal data collections to investigate the setting needs in the English Foundation Programmes. The data collection procedures ranged

from informal discussions with six coordinators, through telephone calls where five of them thought that developing a checklist would help significantly in selecting and evaluating teaching materials, to conducting brainwriting sessions with students and teachers in the English language programme.

To examine learners' needs, many models and inventories have been created to measure or specify the learning styles of individuals. The very famous ones include Kolb's (1984) model, Riding and Rayner (1998) and others like Myers-Briggs type indicator (MBTI), the Felder-Silverman model, the Dunn & Dunn model and the VAC/ VARK model developed by Neil Fleming. The learners are classified in some of these models and inventories into Visual, Auditory, Kinesthetic and Tactile according to their dominant traits. Miller (2002: 2), using Hickcox's (1995) survey of learning, listed these learning style inventories "into three categories". The first is "instructional and environmental preference as in "the Dunn, Dunn, & Price Learning Style Inventory", the second is "information processing preference" as in "the Kolb Learning Style Inventory" and "personality related preference" such as "the Myers-Briggs Type Indicator". These learning styles have been used for many purposes, as indicated by Ehrman, Leaver and Oxford (2003: 314):

Researchers and practitioners use learning style research with personality and cognitive styles to determine ability, predict performance, and improve classroom teaching and learning (Reiff, 1992; Ehrman, 2001; Ehrman and Oxford, 1995). In recent years, the language-teaching profession has also embraced its interpretation of the multiple intelligences model (Gardner, 1983, 2000) as a learning style model for curriculum and materials development (e.g., Gabala and Lange, 1997; Hatch, 1997). Another well-known model adopted by language teachers is the 4-MAT (McCarthy, 1980), which is based on a combination of the brain hemisphericity metaphor (Torrance et al., 1977) and Kolb's (1984) Jung-based model of cognitive style.

Oxford (2003) differentiates between learning styles and learning strategies, as these two terms may cause confusion due to their goals and usage. She defines learning styles as "the general approaches to learning a language", and "strategies, the specific behaviors or thoughts learners use to enhance their language learning" (Oxford 2003: 1). These learning strategies have many definitions, summarized by Barjesteh et.al. (2014: 70) through different researchers' views:

General tendencies or overall characteristics of the approach that language learner employ, and these particular forms of observable learning behavior is appeared in form of techniques (Stern, 1983); “techniques, approaches, or deliberate actions that are employed by students to facilitate the learning, retrieving of both” linguistic and content area information”. (Chamot, 1987); “strategies and techniques that promote the development of the language system and have direct effect on language learning” (Rubin, 1987) and “Behaviors or actions which are taken by learners to make language learning more successful, selfdirected, and enjoyable (Oxford, 1989).

Usefulness of strategies are explained also by Ehrman et.al. (2003: 315):

A strategy is useful under these conditions: (a) the strategy relates well to the L2 task at hand, (b) the strategy fits the particular student’s learning style preferences to one degree or another, and (c) the student employs the strategy effectively and links it with other relevant strategies. Strategies that fulfill these conditions “make learning easier, faster, more enjoyable, more self-directed, more effective, and more transferable to new situations” (Oxford, 1990, p. 8) and enable more independent, autonomous, lifelong learning (Allwright, 1990; Little, 1991).

Despite the spread and extensive use of learning styles inventories in studying and identifying stakeholders’ needs, some researchers prefer to use other means in assessing their learners’ needs. Needs analysis studies that depended on learning styles and strategies inventories are criticized by many researchers including Coffield et.al. (2004), Cassidy (2004), Menaker & Coleman (2007), Jensen (2010) and Rohrer & Pashler (2012). As a result, and because of the dissatisfaction with needs analysis inventories and surveys, an idea-generation instrument was selected for this study in order to specify the categories and items in the teaching materials evaluation checklist regarding learners and teachers’ needs. The innovative instrument used to collect the data about teachers and students’ needs is brainwriting, which is a tool that is widely used in design studies. This instrument is called the ‘6-3-5 brainwriting method’ which is one of the techniques used to collect 3 ideas from a group of 6 participants within six rounds where each round lasts for five minutes. Brainwriting is “the silent, written generation of ideas by a group of people” and “brainstorming, in contrast is the oral generation of ideas by a group of people” (VanGundy 1984: 68).

Brainwriting is one of the idea-generation tools in products development. It can be conducted in a short period of time, it is easy to construct and doesn’t require an expert

facilitator, so any person who wants to develop an evaluation checklist or any other instrument can use it to get plenty of ideas from the involved stakeholders and their needs as it is considered one of the user-centered techniques. Brainwriting “involves silently sharing written ideas in groups” and “relative to brainstorming, brainwriting potentially minimizes the effect of status differentials, dysfunctional interpersonal conflicts, and domination by one or two group members, pressure to conform to group norms, and digressions from the focal topic” (VanGundy 1983, cited in Heslin 2009: 131). In spite of its usefulness, brainwriting can be time consuming in the analysis phase especially if used without the assistance of computer software for data analysis. An alternative instrument can be free listing in which the repetition of the ideas is avoided, and at the same time the stakeholder needs are assessed using a practical and effective instrument.

4.4.3 Needs Analysis Instruments: Students Brainwriting Sessions

In May 2015, permission from the registration office was obtained to get students lists for six groups from three proficiency levels (a, b and c that represent elementary, pre-intermediate and intermediate) to participate in the brainwriting sessions. Also verbal consents were obtained from six teachers to participate in the teachers’ sessions. Later, formal brainwriting sessions were conducted with six teachers and 24 students from different proficiency levels. The needs assessment was required at this stage to obtain an idea about the content, the activities, the tasks and the skills preferred by the teachers and the students in the teaching materials used in the English language programme. The results of the needs assessment were used to design the context-based prototype as a part of the teaching materials evaluation checklist evolutionary prototype.

The six groups that were visited included students from upper-intermediate to beginners’ level. They were provided with an explanation of the research purpose and the instrument used to collect data. Students were asked to participate according to their interest and their willingness to attend the brainwriting sessions. The students in those groups had been in college for two semesters, so they were capable of relating their answers to their experiences. The students who agreed to participate were contacted later and the timing and places of the sessions were specified. The participants were informed of all the procedures required in conducting the brainwriting sessions. The sessions were conducted in single or several groups in accordance with the students’ timetables and their free time slots.

During the students' sessions, it was noticed that higher level groups were more organized and managed to finish sessions appropriately, whereas some lower groups needed more time. Two of the lower level groups were excluded as the participants were unable to adhere and follow the sessions' instruction. The remaining four groups represent three proficiency levels (elementary, intermediate and upper intermediate). The questions were translated into Arabic to make sure that students will understand them to be able to answer the questions correctly. The majority of students asked for permission to write their ideas in Arabic as this would enable them to express their ideas more clearly and efficiently. Ideas from four groups of students were later translated into English in preparation for thematic analysis of the data collected during the different sessions.

Purposeful sampling that was based on representativeness and convenience for the participants was used. First purposive sampling was used when students were selected from different groups to represent different proficiency levels. Second convenience sampling based on (Teddlie, 2007) was used and six students from each group were asked to participate based on their interest and convenience of time and place.

4.4.4 Needs Analysis- Considering the Students Data

Thematic analysis was used to analyse students' data. This is because "a thematic analysis process analyses the data without engaging pre-existing themes" and "each statement or idea contributes towards understanding the issues, which leads to an appreciation of the whole picture" (Alhojailan 2012: 14). Also within this technique, "the process of coding is part of analysis" (Miles & Huberman 1994, cited in Braun & Clarke 2006: 18) and this can help to specify and define the themes and the categories that are going to be part of the teaching materials evaluation checklist.

The categories selected were based on Braun & Clarke's (2006) proposal of doing thematic analysis. This method was used to analyse the huge amount of data collected through the sessions with the 24 students within the four groups in the English Foundation Programme. Within the analysis process, some repeated themes were deleted, similar ones grouped together and other themes that did not belong to the questions asked were discarded. Braun & Clarke (2006: 20) explain the process as following:

Some initial codes may go on to form main themes, whereas others may form sub-themes, and others still may be discarded. At this stage you may also have a set of codes that do not seem to belong anywhere, and it is perfectly acceptable to create a 'theme' called miscellaneous to

house the codes – possibly temporarily - that do not seem to fit into your main themes.

The purposes of the students' questions were to know the students' goals in learning English, their preferred language items as well as their study habits and strategies and techniques. (See brainwriting data collection sheet, appendix D2). Students' purposes for learning English varied from future aims, such as getting a job, to the need to communicate their ideas and communicating with other cultures. Their preferred language content included the focus on the main four skills and other language sub-skills such as grammar, vocabulary and pronunciation. Study habits, defined by Credé & Kuncel, (2008: 429) as "sound study routines, including, but not restricted to, frequency of studying sessions, review of material, self-testing, rehearsal of learned material, and studying in a conducive environment.", were also explored through students' ideas. Their routines revealed the amount of time they spend using their mobile phones and the internet as means for learning. They also still have faith in the usefulness of rote learning and memorizing rules and vocabulary for better language acquisition. Also, they mentioned their preference to start from simple to difficult when they study for their tests and exams. This also showed that they are exam-oriented: they extensively depend on learning the language through studying for exams rather than practicing.

Learning strategies and styles were also investigated through the brainwriting sessions. The analysis of brainwriting ideas regarding strategies, using thematic analysis revealed their emphasis on using drawings, writings and pictures to learn the language faster. Most of the ideas mentioned by the students in the different sessions were also reflected in other practical studies. For example, Griffiths (2008) results on learners' needs revealed that they prefer to learn through the same procedures as learning (1) by hearing language spoken(2) by interacting with others, (3) by memorizing vocabulary(4) by repeating the language many times (5) by learning how language functions (e.g. requesting or complaining) (6) by learning the language related to particular situations (7) by being active in a pleasant environment in a natural environment rather than in a classroom (8) by memorizing grammar rules (9) by using only the target language and (10) by translating to or from my first language from a teacher who is silent as much as possible Griffiths (2008: 261). Also, some classroom observations reinforce students' results in this study as appeared in Kikuchi (2005) such as students interest in talking freely in the classroom, mixing fun with learning, same gender pair and group work, using translation

activities and using various resources as videos, TV, radio and newspapers in the classroom. After the analysis of students' data, the emerged themes mentioned above were used as the items for the students' needs in the teaching materials evaluation checklist in the setting-based prototype. Teachers' brainwriting procedures are described next.

4.4.5 Needs Analysis Instruments: Teachers Brainwriting Sessions

The writers of the textbooks and the educational institutions have to consider teachers' views and suggestions, alongside those of students, policy makers and authorities in materials development and evaluation as teachers' role in materials selection and evaluation is essential in any English language programme. As a result, teachers' views on teaching materials was considered when developing the evaluation instrument.

Teachers' sessions also used brainwriting for data collection. This instrument proved its effectiveness in detecting the stakeholders' needs. Needs here are linked to the textbooks content rather than teachers' interests. Involving a huge amount of work, the results of the sessions helped in clearly defining the materials users' needs and views on teaching materials quality and preferred content and skills.

The first step in the brainwriting sessions was to get consents from the teachers in the English Foundation Programme. After contacting the teachers, consent from six teachers from the College of Applied Sciences in Salalah, was obtained verbally to participate in the brainwriting sessions. One day prior to data collection, an email was sent to the teachers with the instructions and the questions for the brainwriting sessions. The data were collected from the teachers over two sessions on May, 20th & 21st 2015.

Before the sessions, the teachers were informed about the research purposes and the length and time of the sessions as well as the questions through sent emails. Six teachers participated in the brainwriting sessions. They were two females and four males from different countries, ages and experiences. Actually, "the reason for inviting different competences" to the sessions "was to stimulate creativity through the presence of different professional backgrounds, knowledge and experience" (Aschehoug, & Boks 2011: 5).

The teachers were also selected based on purposive sampling which "involve(s) selecting certain units or cases based on a specific purpose rather than randomly" (Tashakkori & Teddlie 2003: 713, cited in Teddlie 2007: 80). The teachers were teaching different

proficiency levels, so they were familiar with materials and it was easy for them to answer the questions by stating their ideas through the brainwriting sessions. This selection of a heterogeneous purposive sampling technique based on “selecting candidates across a broad spectrum relating to the topic of study” (Etikan et.al. 2016: 3) helped to get many ideas that were useful for developing the teaching materials evaluation checklist.

4.4.6 Needs Analysis- Considering Teachers Data

Instructions were sent to the teachers before the sessions along with the brainwriting questions (Appendix D3). Teachers’ questions were about the characteristics of an appropriate English language textbook, the content and the tasks that they would recommend for “good” teaching materials and textbooks. They were also asked about the methods they use to judge the effectiveness of English language textbooks (Appendix D1). After each brainwriting session, the teachers placed similar themes together through affinity diagramming which “is organizing qualitative data into related groups” (Wilson 2013: 34). They did that through the sticky notes that they used to write their ideas on (session pictures are provided in Appendix D4). One of the drawbacks of brainwriting is the participants’ inexperience of grouping the items through affinity diagramming. The amount of ideas presented was huge and some of them were repeated. As a result, a thematic analysis was conducted later to summarize the copious number of ideas into main categories and to specify the items that can be used in the checklist. This technique was used because “thematic analysis is a method for identifying, analysing, and reporting patterns (themes) within data” and “it minimally organises and describes” the collected “data set in (rich) detail” (Braun & Clarke 2006: 6). Thematic analysis “also often goes further than this” as it “interprets various aspects of the research topic” (Boyatzis 1998, cited in Braun & Clarke 2006: 6). Basically, this instrument “can be applied within a range of theoretical frameworks, from essentialist to constructionist” (Clarke & Braun 2013: 120). It is also similar to affinity diagramming that is usually used with brainwriting sessions and which will enable further summarizing of teachers and students’ data. In addition, thematic analysis allows for more flexibility “as this stage of the research is an initial step”, and the researcher “need not subscribe to the implicit theoretical commitments of grounded theory” as the aim is not “to produce a fully worked-up grounded-theory analysis” (Braun & Clarke 2006: 8), but rather to specify the most conspicuous themes that can form the categories and items for the evaluation checklist development. Also, “a theme might be given considerable space in some data items, and

little or none in others, or it might appear in relatively little of the data set” therefore, the “researcher judgement is necessary to determine what a theme is” (Braun & Clarke 2006: 10). The processes of thematic analysis included “Familiarisation with the data, Coding, Searching for themes, reviewing themes, Defining and naming themes and Writing up” (Clarke & Braun 2013: 121-122). This method of data analysis uses three approaches to coding data: “theory driven coding, research driven coding and data driven coding” where data driven coding “involves inductive code development based on the data collected in the study” (Kawulich 2004: 99). The last type of coding was used in specifying the themes in the data collected from the teachers’ brainwriting sessions.

The results of the analysis were sent to the participated teachers three times, after each process of coding, to make sure that all their important ideas were there and that the categories and the items that were developed were properly placed. The final categories and items were also sent, to check for any inaccuracies or overlooked ideas. This is because, “when gathering sub-themes to obtain a comprehensive view of the information, it is easy to see a pattern emerging” and “when patterns emerge it is best to obtain feedback from the informants about them” (Aronson 1995: 3). These are also the strategies followed to “attain trustworthiness” which include “peer debriefing, prolonged engagement and member checks...” (Morse et.al. 2002: 15).

There were many innovative ideas raised by the teachers. For example, teachers mentioned CLIL (content and language integration in language learning) as one of the issues that ought to be considered when developing textbooks or teaching materials. “CLIL is used as a brand name to add value to General English coursebooks” (Banegas 2014). In addition “the current ELT global coursebook market has embraced it as an innovative component for general English coursebooks for EFL contexts” (Banegas 2014: 345). The teachers also mentioned critical thinking skills as well as study skills, which they think that textbooks have to integrate through their tasks and within their content. Also the use of gamification is mentioned which “proposes the use of game-like rule systems, player experiences and cultural roles to shape learners’ behavior” (Lee & Hammer 2011: 3). According to Lee and Hammer, there are “three major areas in which gamification can serve as an intervention” cognitive, emotional, social” (Lee & Hammer 2011:3). Some teachers therefore thought that incorporating such games in the teaching materials will help the students to learn the language. Many of the teachers’ ideas were incorporated into the teaching materials evaluation checklist as they were reflected in

their brainwriting sessions (see appendix 5 for a full summary). Teachers' thoughts and views do match other empirical studies as they almost share the same topics of interests and concerns. The matching between the results of the teachers' data and other studies done by different researchers can be attributed to the teachers' personal experiences in the classroom as well as their interest in the problems of their students more than their own concerns. This is exemplified by a whole chapter in McGrath (2013, 105-126) where the results of the empirical studies on how teachers evaluate textbooks revealed teachers' emphasis on making sure that the materials fulfil the needs of their students as a first priority for them. As McGrath, 2013, p.118) explains "teachers were perhaps predictably more influenced by learners factors...and practical considerations" McGrath, 2013, p.118). Another study conducted by Richards(1993) indicated that teachers do pay more attention to the materials that best suit their learners' needs such as matching the course objective, being "easy to use", fitting "class needs", being "culturally appropriate", "can be used with classes of mixed ability" and teaching points are easy..." Richards (1993, p.11). Other studies by Sercu, Mendez Garcia & Castro Prieto (2004) and Alamri (2008) in McGrath (2013) shared the same teachers' conceptions as they appeared in this study where they do focus more on the students' lacks and needs. In that context, it is revealed by Kayapinar that "the general conceptions of the teachers suggest that coursebooks should be developed and used to meet the needs of the learners in the national context." Kayapinar (2009: 69). Accordingly, the analysis of data for both students and teachers signposted some related themes which all emphasize the importance of considering the views and opinions of different stakeholders in the English Foundation Programmes.

4.4.7 Institutional Needs Short Survey

The short survey was sent to a key informant in the Ministry of Higher Education in Oman. Within her responses, she stated that "there is a need to update these books and provide ones that present innovative methodologies and ground-breaking strategies for teaching in order to address the need of the new generations." This goal can only be achieved when there is constant evaluation of the teaching materials in the English Foundation Programmes that involves "teachers' Feedback, students' feedback and comparing students' performance using standardized international tests. According to her point of view, "students' proficiency levels, students' majors and degree programmes, Foundation Prgrammes National standards and the price of the textbooks" are important when selecting general English textbooks. Thus, it is important to find a tool that enable

educational institutions to achieve their goals as their roles in quality management are very important in Oman. Al-Issa (2005: 110) state in a study conducted in a similar context that “the success of the Omani higher education in delivering quality (language) education is largely, if not entirely based upon the efficiency of the school system”. Most of the categories were incorporated in the checklist first prototype, but only two items are remained in the checklist after its review by different experts and expected users. These are: (1) materials should consider the Ministry of higher Education standards when selecting and evaluating teaching materials and (2) materials should provide methods for “cross check of the students’ performance using standardized international tests”.

The defined themes from teachers’ data along with the ones from students’ data, as well as the short survey of the institutional needs, were used to develop setting needs prototype. The result of this cycle of data collection and analysis is the context-based prototype that is merged later with the theoretical-based prototype to form the first evolutionary prototype of the checklist for teaching materials evaluation in the English Foundation Programmes. These stages which included literature review and data collection are part of the design-based research methodology phases, as explained by Sahasrabudhe et. al. (2013: 3):

Each iteration of modifying the intervention is termed as a research cycle. The important aspect of DBR is that the outcome/s of every research cycle is used as input for the next research cycle ...This helps in augmenting the intervention on the basis of the ‘failures’ in the earlier research cycles. The cycles conclude after a particular version of the intervention shows desired results. In DBR, the conclusions of this process not only have the detailed log of the chronological development of the intervention but also the documentation of the problems recorded in the earlier cycles (along with the steps taken to address them).

4.4.8 Summary

By the end of this stage, two important constructions have been established through this study. The first is the conceptual framework that can be used to explain the teaching materials evaluation checklist sources, basis and development. The second is a checklist for teaching materials selection and evaluation in the Colleges of Applied Sciences English Foundation Programme. Following the construction of the first teaching materials evaluation checklist prototype, the next step was started to test the checklist

through several cycles of formative review, preceded by a screening of the checklists using several references from the literature.

This prototype went through many revisions attempting to avoid overlapping between the main categories and items. This stage was probably the most complicated process, as some concepts of educational terms and definitions are strongly intersected. A satisfactory shape for the prototype was finally accomplished to be tested through the several iterative cycles of formative review by the experts and the potential users. The evolutionary prototyping is used to “continually refine” (Nieveen, 2009: 90) the checklist through formative review cycles (one-to-one evaluation, expert review, small group review and field testing).

As a start of the formative review which included several cycles in this study, the checklist was revised against four of the models. Using developer screening the checklist was reviewed through comparing the developed checklist with the four models: Tomlinson (2013), Bichelmeyer (2003), Stufflebeam (2000) and Wilson (2013). During this stage, efforts were made to ensure that the items in the checklist prototype were compatible with the results extracted from research and from the setting needs data collection instruments. For example, issues like asking only one question in each criterion, the possible sources for content and the general format for the checklist were considered during this cycle.

Most researchers use more than one technique of ‘formative evaluation’ or ‘formative review’ ,as it is called in this study, such as developer screening, one-to-one reviews, expert reviews, small group reviews, field testing and even summative reviews of their developed tools and products. Formative review can be defined in that sense as a multi-cycles method with three distinguished features “ongoing” process that “involves assessment” which “seeks specific information as well as judgements” (Beyer 1995: 7). Van den Akker (1999) explains the importance and the role of formative review in design-based research studies indicating that “the main reason for this central role is that” formative review “provides the information that feeds the optimization of the intervention” which helps “developers during the subsequent loops of a design and development trajectory. It is most useful when fully integrated in a cycle of analysis, design, evaluation, revision” Van den Akker (1999: 10).

In this chapter, the categories for research-based categories were specified from literature and the practical needs categories were identified through brainwriting sessions and a

short survey. Based on these procedures, the checklist prototype was developed and refined through developer screening. The next stage is the other formative review cycles, an essential part of most of the design-based research studies as well as the assessment method of the teaching materials evaluation checklist developed in this study.

Chapter 5 The Checklist Review

5.1 Introduction

Evaluation and reflection phase follows the design and construction phase. Once the instrument or the prototype is constructed, it has to be tested with the intended users and participants through a systematic method such as formative evaluation. Indeed, formative evaluation as a DBR method “enables researchers to explicitly study the complexity of implementation projects and suggest ways to answer questions about context, adaptations and response to change” (Stetler et. al. 2006: 1). This method is widely used in many fields to develop products, systems, models and processes. But in education, it is mostly connected with programme and course evaluation. Baker & Alkin (1973: 389) advised that a “developer should engage to a greater extent in formative evaluation of all products” because “formative evaluation data would provide information to developers that would allow them to modify and improve their products before they are distributed”. The planning of the formative evaluation is very important so that the “problems encountered during data collection” can be “anticipated in planning evaluations and dealt with by appropriate procedures in the conduct of studies. In practice, however, evaluations are rarely designed and executed perfectly” (Burstein et.al. 1985: 68). To ensure that the data will be collected accurately and efficiently, and according to Reigeluth & Frick (1999) some useful techniques for collecting the formative data include screening, pilots and tryouts. These techniques should be prepared in advance and scheduled professionally. Nieveen & Folmer (2013) introduces a detailed definition that is “based on a comparison and synthesis of the definitions of various scholars in the field of formative evaluation” such as (Brinkerhoff et.al. 1983, Flagg 1990, Scriven 1967 and Tessmer 1993) and which defines “formative evaluation in the context of design research as *a systematically performed activity (including research design, data collection, data analysis, reporting) aiming at quality improvement of a prototypical intervention and its accompanying design principles*”[italic in source] (Nieveen & Folmer (2013: 158). In a shorter but comprehensive definition, formative evaluation is defined by Dick (1980: 3) as “the process of collecting data about a product during its development”. Throughout the process of this study, it was noticed that there are misconceptions around the term “formative evaluation”. The first misconception is its confusion with the formative assessment of students’ learning during the course of their study. Despite the fact that formative evaluation has other uses and purposes which are totally different from

formative assessment in the classroom settings, some users are still confused about the two terms. The second misconception could happen because of the title of the instrument developed within this study -teaching materials evaluation checklist- which can be confused with its validation process of “formative evaluation”. Looking at both terms “formative evaluation” and “formative review”, it can be inferred that what happens through using this method in design-based research is more a review process rather than an evaluation process. Each prototype is reviewed by experts and other users such as teachers to identify potential problems, not merely evaluated to identify weaknesses and strengths and make a judgement. Third, the term itself has kept its label since its coinage by Michael Scriven in 1976, despite the growth of its purposes and uses in different types of research, including educational research. To renovate its use as a systematic method in design-based research, a new and clear label has to be used. By clarifying such terms, the “iterative” nature of both formative evaluation and design-based research can be easily recognized by researchers and different participants in any study that uses both DBR as methodology and “formative evaluation” as a research method with its different instruments and cycles. In this thesis, the term “formative review” will be used as the method for data collection to assess and validate the developed checklist for teaching materials evaluation in the English Foundation Programmes.

5.2 Formative Review

George & Cowan (1999) specified several reasons which initiated and promoted the use of formative review in education. These reasons included the onset of accountability which was “more and more demanded of professionals” the “habit of self-critical reflection” that “moves the practitioner beyond Kolb’s cycle, which concentrates on the development of abilities, and where the question ‘how do I do it?’ was changed into ‘how well do I do it?’. Finally, “the introduction of a system of accreditation for teachers in higher education” as well as the “many advances in learning and teaching contexts that require the use of evaluation” on a regular basis (George & Cowan 1999: 2-3).

Formative review “has traditionally involved four major stages: expert review, one-to-one evaluation, small group evaluation, and field test” (Lake & Tessmer 1997: 5). The developed checklist through this study was tested using the four cycles of formative evaluation: one –to-one evaluation as well as “expert appraisal, micro-evaluation and tryout or the field test” (Nieveen 2007: 95-96). Prior to these stages, and through the developer screening cycle, the prototype of the checklist was checked against four sources

on how to develop evaluative checklists, detailed in chapter four, which included Tomlinson (2013), Wilson (2013), Bichelmeyer (2003), and Stufflebeam (2000). Then, the checklist prototype was sent to four experts in teaching materials development and evaluation: Prof. Brian Tomlinson from the United Kingdom, Dr. Saleh Al-Busaidi from Oman, Prof. Jaykaran Mukundan and Dr. Vahid Nimehchisalem from Malaysia. The checklist was also revised by three teachers in one-to-one reviews. This was followed by the review of a thorough assessment in small groups (with 2 experts and 6 teachers) and lastly, a field test using the checklist prototype in evaluating teaching materials with teachers and coordinators in the six Colleges of Applied Sciences in Oman. The first rotation started with the expert review, because the experts' role is essential (besides triangulation of data and sources) for "scientific rigor" of the checklist in its early and final stages as the researchers "may easily become too 'attached' to their prototype which could lead to a less objective view toward problems and comments from the respondents" (Nieveen 2007: 99). Each stage of formative review helped the researcher to become detached from the developed checklist and to discover more of its problems, weaknesses as well as its strengths.

5.3 Formative Review Stages

The formative review of the teaching materials evaluation checklist is based on Tessmer (1993) where the data is collected through four cycles. Each cycle has its own data collection instruments. For example, in expert appraisal for the checklist review, two instruments were used: four open-ended surveys and in-text annotated analysis. In Tessmer's model, the expert review and the one-to-one review can be conducted at the same period of time. The results of those two types of reviews were then used to amend the product or the instrument designed: the teaching materials evaluation checklist for English Foundation Programmes. The results of the two cycles can also be compared in order to identify the views of both experts and teachers regarding the teaching materials evaluation checklist. These stages are clearly depicted in Figure 7.

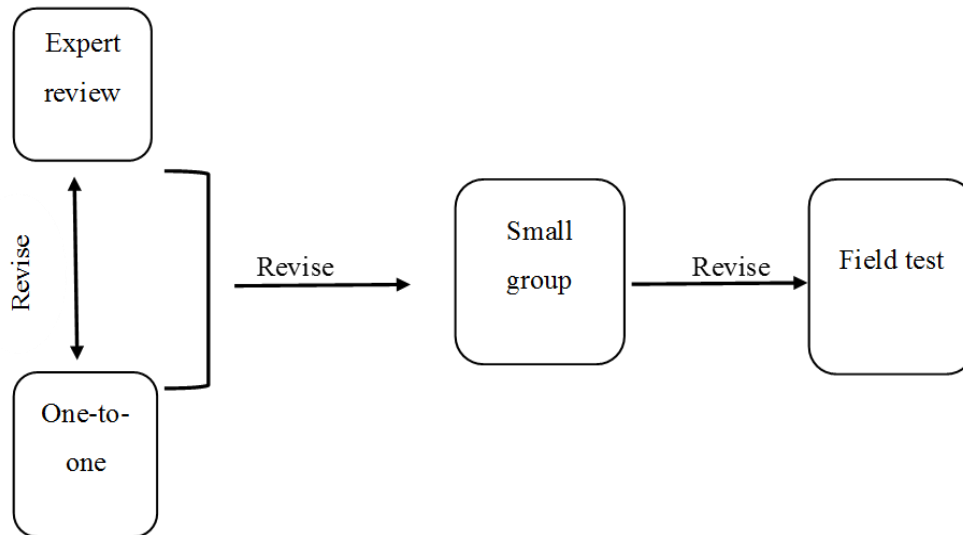


Figure (7) General Sequence of Formative Review Techniques (from Tessmer 1993)

5.4 Designing the Formative Review Instruments

The design of the review instruments for each cycle, went through different stages. First, each cycle objectives were specified. For example, the expert review main goal was to test the internal validity of the checklist, whereas the one-to-one review aim was to get a first impression from the participants when dealing with the checklist for the first time. It was intended also to discover the participants' general view on its appeal, comprehensiveness and usability. The objectives of the small group review were to discover some of the problems regarding the use of the checklist in an environment similar to the real context in preparation for the field test cycle, as well as the checklist's effectiveness, practicality and appeal. Field testing aims were to discover usability problems and to observe how the checklist can be used in real settings, so the instruments were designed to suit the objectives of each cycle. For each of the four formative review cycles at least two instruments were used to collect data. The instruments for data collection were peer reviewed by some colleagues in the College of Applied Sciences in Salalah to check language and comprehension problems.

5.5 Participants' Sampling

The number of participants in a formative review is usually small because the data will be collected through different cycles and the number of participants has to be manageable especially if the study is conducted by a single researcher. In fact, the formative review of the instrument and any other intervention can be considered usability tests according to Petrie & Bevan (2009: 22) where sample size requirements for a particular desired

percentage of problems can be estimated by calculating the probability of finding problems, either based on previous similar usability evaluation, or from initial results of an ongoing study. A recent survey by Hwang & Salvendy (2007) found probabilities in the range 0.08 to 0.42. This would correspond to evaluating the checklist with between 3 and 19 participants to find 80% of the problems, or between 4 and 28 participants to find 90% of the problems. So “the number of participants used is based on how many participants are needed to reasonably determine” the problems that were not discovered previously (Medlock et. al 2002: 1). The general purposes of the formative review were basically three as illustrated in Table 7. The table also shows the samples that range between 3 to 8 participants and the description and time for each of the three types of reviews.

Purpose	Description	When in Design Cycle	Typical Sample Size (per group)	Considerations
Early formative evaluations				
Exploratory	High level test of users performing tasks	Conceptual design	5-8	Simulate early concepts, for example with very low fidelity paper prototypes.
Diagnostic	Give representative users real tasks to perform	Iterative throughout the design cycle	5-8	Early designs or computer simulations. Used to identify usability problems.
Comparison	Identify strengths and weaknesses of an existing design	Early in design	5-8	Can be combined with benchmarking.

Table (7) Purposes of user-based evaluation and the recommended number of participants by Petrie & Bevan (2009).

The formative reviews conducted on the checklist were diagnostic in nature with exploratory and comparison instances that were based on the cycles goals and the analysis of the participants feedback. The goal was to discover the problems of the checklist in its developmental stages. Purposeful sampling was used as the participants were selected based on their different backgrounds and experiences as well as the proficiency levels they were teaching in the English Foundation Programme during that semester. The first procedure for collecting data using formative review was through preparing a good plan for the data collection in the different stages that include the number of participants and the arrangements for different sessions and reviews.

5.6 Data Collection Arrangements

To prepare for data collection, first a schedule was created to specify the number of participants and their institutions. Second, the instruments for each cycle of formative review were designed. Third, the participants were contacted through emails and through their institutions. Then, as soon as the permissions and consents were received verbally or through emails, the time and the places were arranged in collaboration with the participants in different institutions. Each cycle required different preparations in terms of the locations and facilities. For example, the small group review for teachers was conducted in one of the English Department computer labs as it required power point presentation before the beginning of the session. Also, the field testing sessions had to be conducted face to face or online in the six Colleges of Applied Sciences.

5.7 Data Collection Processes of the Formative Reviews

The data collection processes varied in terms of time, the break between each cycle and the amendments performed on the checklist after each cycle. For example, the expert appraisal and the one-to-one review took one month to finish, as one of the experts was really busy and it took him more than the two weeks specified to return his feedback. Other cycles also needed a longer time than expected due to problems of contacting participants, scheduling the sessions and making the appropriate modifications on the checklist after each cycle.

5.8 The Expert Review of Prototype 1: Sampling and Participants

The search for experts in materials development and evaluation took some time as some experts from Oman, the United Kingdom and the United States were unable to participate because of their busy schedules. Of the many experts contacted, four experts agreed to be

part of the study. Their experience was basically in teaching materials, their development and evaluation. Tessmer (1993: 51) calls these experts “subject matter expert (s)” which refers to “someone who has acquired current and thorough knowledge” about the investigated topic. The subject matter experts or the SME “would have both theoretical and practical knowledge.” Therefore, the experts who were contacted for this study had both aspects as they are academic researchers, teaching materials developers and evaluators. The reference to these experts within this study will be as follows: Prof. Jayakaran Mukundan: Expert 1, Dr. Saleh Al Busaidi: Expert 2, Dr. Vahid Nimehchisalem: Expert3 and Prof. Brian Tomlinson - Expert 4. The four experts’ short bios are available in (appendix A). They were contacted through emails and were asked for their consent to participate in the study and upon their agreement, the instruments for data collection along with the developed checklist were sent to them.

5.8.1 The Instruments Used in Data Collection for Expert Reviews

The experts were sent a package that included a short summary of the study, the survey questions and the developed checklist. They were asked to answer the questions as well as to write their comments directly on the checklist. The instrument used was a four open-ended questionnaire which included the following questions:

1. What are your suggestions regarding the content of the checklist in terms of its:
 - a. Inclusiveness of all the necessary items for a “general English” teaching materials evaluation checklist.
 - b. The precision of its words and terms.
2. What are your suggestions regarding structure of the checklist in terms of:
 - a. The grouping and sequencing of the specific items within the main categories.
 - b. The transparency of the checklist’s layout with reference to the main headings and sub-headings, numbering, organizations of items and attractiveness to its prospective users.
3. The reliability of the checklist in terms of its generalizability to other contexts and how it can be adjusted to suit various settings.
4. Further suggestions on:
 - a. Any other sources, categories or items to be added to the checklist?
 - b. Any other methods that can be used to judge or evaluate the checklist?

As can be noticed in the above feedback questionnaire, the purposes of these questions were to check the validity of the checklist in terms of its content precision, inclusiveness, its structure regarding sequencing and layout as well as its generalizability. The experts were also asked to write their annotated comments directly on the checklist, so they could cross out, change or add any category or items with the aim of improving the checklist content and structure. The expert responses were received within one to four weeks from sending the emails. Some completed it in the week in which it was received, others needed more time as they were very busy. The feedback of the experts will be discussed next.

5.8.2 The Experts' Feedback

The experts' feedback was very important at this stage of the checklist development. It was essential to make sure that the content, the structure and the format of the checklist was well developed before the other reviews were conducted. In analyzing the data from experts' review, Tessemmer (1993: 63) referred to three main actions in dealing with their feedback. The first is to reject the "ones that would lead to pointless or impossible revisions". Second, the developer has to "summarize and list the expert's comments" and finally, he/ she can note the "area of agreement and disagreement" within the experts' feedback and comments. These processes will help to "understand what areas may require improvement" and reveal "the types of revisions that should be made." The experts feedback consisted of their answers to the above questions (see their full answers in Appendix G2) and their annotated comments on the checklist itself. Their comments included the following points:

In reply to question 1 about the content of the checklist, concerns about the length and rewording of the checklist were raised by Experts 2, 3 and 4. Some of their comments comprise these points:

The checklist is immensely too long. Teachers will certainly find it time consuming to use! I wonder if a shorter, teacher friendly checklist could be developed based on it. (Expert 2)

Some of the items could be reworded more effectively. Some items look more like notes than clear-cut, unidirectional and terse items. (Expert 3)

a. It's a comprehensive list. I'd be tempted to delete some of the less useful criteria to make the checklist more user friendly.

- b. Sometimes you use the words of experts who haven't used them with precision or clarity. I'd suggest using your own words and being a little more constructively critical of the experts. (Expert 4)

The comments for question two about structure were fewer, as most comments were annotated directly on the checklist. The overlapping between some items in the main sections or categories was referred to by Expert 3: "the sections and sub-sections look inclusive but since they come from different sources, the developer should make sure they do not overlap".

For the third question about the generalizability, the comments were about the importance of simplicity of words (Expert 2) and "avoiding items that look into more than one particular subject matter" (Expert 3). Expert 4 was more concerned about the division of the checklist main categories:

I'd divide the criteria into universal criteria and local criteria. The universal criteria apply to any learner anywhere and are therefore generalizable to other contexts. The local criteria are specific to the particular context of your evaluation and are not transferable to other contexts without modification. Universal criteria derive from principles of language learning. Local criteria derive from a profile of the particular learning context. (Expert 4)

Question 4 comments were also included within the checklist. Expert 4 was most concerned about the mixing of the analysis and evaluation items in the checklist. His comments included differentiation between analysis and evaluation in details:

The Impressionistic evaluation checklist is a mixture of analysis (i.e. 1 and 2) and evaluation (i.e. 3 and 4). Also I don't understand how to use the three availability columns in answer to suitability questions. Your checklist would be much more user friendly and reliable if you separated analysis from evaluation, if you used criteria to phrase questions, if your analysis questions were Yes/No questions and if your evaluation questions were answered on a scale from 1-5...Remember that with an analysis you are finding out what the materials consist of and what they ask the learners to do. With an evaluation you are predicting the likely effects of the materials on their users. (Expert 4)

These were some extracts on the invaluable feedback from the experts on the teaching materials evaluation checklist first prototype. More of their comments that were used in improving the checklist and the most important changes that were considered are detailed in the next sections.

5.8.3 The Changes Made to the Checklist Prototype 1 (Experts Reviews)

Following the experts' answers to the four questions, their annotations on the checklist and their recommendations were utilized to make the checklist's content, organization and layout more accurate, effective and practical. Four main amendments were implemented: deleting some items, merging some, adding some and modifying others. There were items which were deleted because they were repeated in other sections, or which may confuse the evaluators. For example, the column that referred to the sources of the items in the checklist was deleted as one of the experts suggested that it is not useful to include it in the checklist. That column can be included in a guide or a website, which can be developed later, on the heuristics of designing and using the checklist. Another example is "multi-level" as one of the experts suggested the teacher may not "understand what you mean by multi-level". Some of the other items that were removed from the checklist are basically in the first three categories of the checklist. For example, the item "materials should maximize learning potential by encouraging intellectual, aesthetic, and emotional involvement which stimulates both right and left brain" in Second Language Acquisition section, was removed, though a few thought it is a good standard, it caused some problems for others, and this may complicate the teaching materials evaluation process (see the checklist prototype 1 in Appendix L1). It contained complex terms, which makes it difficult for the evaluator to conduct the materials evaluation. As Expert (1) suggested, it is an area "where the textbook is incapable of doing very much" in his annotated comments in the checklist.

The other items within the main category "literature review" which was changed later into "research", were mostly in sections two and three. They were omitted because of two main reasons: repetition and inappropriateness for the materials evaluation checklist functioning. Items that were removed included (a, b, d, and h see appendix L1) in section two (teaching principles and pedagogical factors). Some of the experts' annotated comments are illustrated below:

"There are too many factors for this to be a criterion. Phrased as a question it would be unanswerable". (Expert 4 opinion about item /a/ in section 2)

"You should think of explicit and clear-cut criteria. This is at the moment open and vague. (Expert 3 opinion about item /a/ in section 2)

"But how can the most favourable sequencing be determined? It will be different for each learner. Also there is a false assumption that the

sequence of teaching will necessarily be the sequence of learning.”
(Expert 4 opinion about item /d/ in section 2)

Other examples of overlapping between items is /b/ and /d/ in section two also, where two experts' notes were about being careful about overlapping: “not clear the difference between this and b” (Expert1) and “watch for overlaps” (Expert 3). There are also items that were moved in the setting needs section under the category “teachers’ needs”. These items were moved from /f/ to /a/ because they were related to the teachers’ general view about English language materials. The suggested items were the following:

- Can be modified and edited through the availability of soft version of materials
- Whole package with supplementary materials and CDs
- Teachers friendly
- Clear instructions and easy to be taught
- Publisher available for questions and feedback

Some of the items were modified, such as the title of the checklist, which was changed from The Teaching Materials Evaluation Checklist for General English (TMEGE Checklist) to Teaching Materials Evaluation Checklist for English Language Programmes (TMEC for ELP) as two of the experts expressed their concerns about its wording, such as the use of “General English” (Expert 1). The rating scale of the checklist was also amended to make it easier for users as there were some of the items that were merged in the setting needs part under the category “teachers’ needs”. Expert (4) proposed that “there are two different criteria embedded in the one criterion” so “the appropriate content and language items for listening and speaking” became “the appropriate content for listening and speaking” as the word “content” is representative of both criteria. Some of the phrases were also abbreviated for ease of use, such as “practical considerations”, which was replaced by “practicality”. Finally, an Excel sheet was used instead of a Word document to make the rating scale more accurate when calculating the checklist scores during evaluation. This procedure required the assistance of an expert in designing the checklist prototype Excel sheet. He helped the researcher in writing the appropriate codes for the rating scale and in fixing the bugs that appeared later when the Excel version of the checklist prototype was used by the users in the small group and field testing. One-to-one review sessions are explored and reported next.

5.9 One-To-One Review of the Checklist Prototype 1

5.9.1 Introduction

Before the experts' feedback was received, the one-to-one reviews took place in the College of Applied Sciences in Salalah. Three teachers were contacted to participate in the study. They also had different backgrounds and were teaching three different proficiency levels in the English Foundation Programme. The aim of this sample selection was to know how different teachers would approach and use the checklist. The one-to-one review main purposes were to investigate the checklist "clarity, appeal, errors, practicality and usability" (see Tessmer, 1993: 74-75).

5.9.2 Sampling and Participants

The teachers in one-to-one reviews were asked directly if they were interested in participating in the study. Then, the time of the sessions was allocated for each teacher who had agreed to participate in the sessions. The three teachers varied in their experience, their gender and their nationalities. The first participant (Teacher 1) is a British male teacher who has a post graduate education and 16 years' experience in teaching. The session was conducted on November the 8th 2016 and took about 45 minutes to finish. Teacher 2 is an American female teacher who has a Bachelor degree in Education and 21 years' experience in the field of teaching. Her session was on November the 9th 2016 and lasted for about 32 minutes. The third teacher (Teacher 3) is an Omani male teacher who has a Master's degree in Education and 3 years' experience in teaching. The session was conducted on November the 10th 2016 and it lasted for about 1 hour and 25 minutes.

5.9.3 Procedures and Instruments

The one-to-one protocol used in this review is based on comprehensive procedures recommended by Snyder (2003) and Beyer (1995) as detailed in (Appendix H1). In this protocol, there are three types of data collection: evaluator questions that were asked after the teacher's first reading of the checklist, observation logs for notes, and debriefing questions at the end of the session. The researcher questions were mainly about the checklist clarity, completeness and ease of use. The recommendations made by the three teachers were mostly incorporated and some changes were made in the teaching materials evaluation checklist, which are demonstrated below.

5.9.4 Changes Made In the Checklist Prototype 1 (One-To-One Reviews)

To deal with data collected from this stage, certain techniques were used. Tessmer (1993: 89) suggests using a “data sheet” to combine and group data “by the type of information given” such as “comments, performances” and the “behavior observed”. These processes helped to identify the areas that needed amendments and revisions. Considering the three teachers’ data, it was noticed that in the general comments, the three teachers agreed that the checklist was comprehensive and inclusive. Teacher 1 suggested adding a column at the end of the checklist for the evaluator’s notes so that “the evaluator can express his/her opinion”. In response to clarity questions, Teacher 1 said that the checklist was “easy to understand”. Some of the other comments were about the use and the connotation of some words such as “rural”, “travelers” and “simple lives” (item i in second language acquisition principles) as the meaning of “simple lives” differs in developing countries and rich ones. For example “farmers in the UK tend to be rich whereas in developing countries they are usually poor” (Teacher 1). Also “travelers” has a different connotation in western countries: in the “UK, for example, it is connected to gypsies” (Teacher 1). Though “the categories are clear” for Teacher 2, he agreed with Teacher 1 that the item “materials should reflect the reality of language use” “needs more explanation” as the “reality of language use varies from one person to another” (Teacher 2). Teacher 3 had the following comments about clarity: the main category A (research-based aspects) could be replaced by curriculum theories or learning theories and though these headings were not changed, the terms describing them were simplified. The same teacher commented also on the “precision in phrases and terminology.” Also, he suggested that item /g/ needed more details and examples on “aesthetic and emotional involvement.” There was also a recommendation to “replace lives with life styles.” Also Teacher 3 pointed out to item /b/ in the category entitled “students’ needs” where “some are specific and some are general” and others that are not clear, such as “understanding key strategies” and repeating and recycling” and his question was “which skills or inputs to repeat”? Teacher 3 comments also included “suggesting guide, manual for use and for explaining some items.”

Among the comments about the checklist’s completeness concerns were raised by Teacher 2 about item /b/ (see Appendix H4 for this teacher feedback) and about the favorable sequencing. In the students’ needs section, two teachers (1& 2) asked for “the

goals of learning English” to be included in details. The three teachers insisted on “listing what you are looking for in all sections”, which means that the column in the checklist “what to look for” was very important for the teachers as a guide in their evaluation of the teaching materials. Unlike the view of Expert (4) who advised the researcher to use her own words, Teacher 3 thought that the use of the exact educational terms in the literature would make the checklist more understandable, as he recommended to use “terms from the literature” for “educators”. He also suggested a “guide/detailed description of the purpose of categories and items used.” Teacher 3 also asked for an “explanation of the process of categories and items selection” and raised inquiries about the “basis of selection? How did you reach this? Do the categories cover all the required areas?” In fact, these questions were also raised by the researcher when thinking about designing a checklist, which indicates the importance of the conceptual framework that was developed and validated by four experts in the first phase of this study. This framework has dual functions: to work as a basis for the checklist development and a guide for the whole research design and processes. It actually answers all the questions raised by Teacher 3.

Regarding the checklist’s ease of use, Teacher 1 thought that the checklist could be difficult for some “especially if the teacher’s major is in different field, and has taken TESOL to be able to teach.” Teacher 2 suggested that the checklist was a tool that “gives you the sense that “I have a job to do” and can be done “thoughtfully within 1-3 hours.” On the other hand, Teacher 3 assumed that “it will take time if good evaluation has to be done” and he suggested that “starting with training sessions will help.” He also recommend a guide or manual for the teachers on how to use the checklist. The teachers also noticed a few typing mistakes that were all corrected in the checklist prototype. The three teachers thought that the checklist was implementable, but if “given a chance” (T1) as “it may help the teachers to be more critical about textbooks.”(T3).

The researcher’s observations throughout the three teachers’ sessions revealed that the checklist is easier for experienced teachers. The teachers who are English native speakers, or who had more experience, seemed to finish the checklist review in less time (32 minutes and 45 minutes) and with less effort, while it took the third teacher one hour and 25 minutes. They had also less misunderstanding and no long pauses. Teacher 2 explained that the pauses were “because the teacher was “comparing” the checklist items with what she sees in the textbooks she is using.” Teacher 3 was a little bit confused and

he explained that during the pauses, he was “attempting to answer that specific item and isolate it from others” for each section. The above comments helped to raise more questions regarding the usability of the checklist among some teachers, native and non-native speakers of English, which will be considered in its revision.

The debriefings at the end of the three sessions elicited few comments as most of the suggestions were expressed through the researcher’s main questions. When asked if they thought that other teachers would be using the checklist, Teacher1 thought that “some would” and “some wouldn’t: it depends on time and payment for the teacher.” Teacher 2 also shared the same opinion with Teacher 1 as she thought that some would be interested “given a chance”, but “some won’t be interested.” The last question was about the things that the teachers learned from the checklist. Their answers were interesting as Teacher 2 answers illustrated: “It did refresh things” from my experience in teaching... “It made me think why you like something in a textbook or why you didn’t like it”. Teacher 3 said that the checklist showed that he needed “to know more about curriculum design.” At the end of the cycle, the researcher made use of all the comments received from both the four experts and the three teachers to improve and revise the prototype of the teaching materials evaluation checklist. The checklist prototype had evolved into a semi-complete version that can be tested for more improvements and adjustments. In spite of the importance of the previous two cycles, the next two cycles (the small group review and the field testing) will be very important to discover more issues about the checklist usability and practicality for its potential users (experts, coordinators and teachers).

5.10 Small Group Review of the Checklist Porotype 2

5.10.1 Procedures

As mentioned in the previous section, the model that is used for data collection in this study is based on Tessmer’s (1993). In view of that, the evaluation of the checklist could move to its third cycle. This cycle consisted of the reviews of a small group of six teachers and four experts according to the prepared plan. Tessmer (1993) advised to select “a group small enough so that (the researcher or product developer) can manage the data analysis” as “open ended questions and debriefing are difficult to manage with large numbers” (Tessmer 1993: 108). Two of the experts, were very busy during the time of conducting these sessions, so they were spared to participate in the field testing cycle. Beyer (1995: 16) suggests that “to secure the most useful feedback, we should ensure that several experts examine the product” so, having experts in the field testing will enable

the researcher to get more feedback on the checklist weaknesses and strengths. Indeed, having the two groups of users (experts and teachers) in each cycle, helped the researcher to make a comparison between the two groups in terms of the checklist practicality, effectiveness and appeal to the expected end users.

5.10.2 The Small Group Review (Experts)

Experts are considered by Beyer (1995) as one of the three types of people who are “essential in the formative evaluation of any educational product”. The other two types are “users and stakeholders”. The experts contacted for this review cycle were considered “knowledgeable about the intended users of the product” (Beyer, 1995: 21) as well as having responsibility for teaching materials development and evaluation in the English Foundation Programmes. The researcher contacted four experts from two higher education institutions, but only two experts were able to participate in the review sessions.

5.10.2.1 Sampling and Participants

In order to reach the experts, the Foundation Programme director was first contacted as he was able to nominate the most appropriate person for the experts’ reviews. The relevant institutions were visited and two teachers were contacted, who were able, as insiders, to direct the researcher to the suitable experts. Then the experts for this cycle of data collection were contacted through emails with a short explanation of the nature of the study and their roles in the small group reviews. Later, the checklist and the feedback questions were sent to them and the time for the session was specified. The two experts were PhD holders with experience that ranged between 10 and 16 years. The feedback from these two experts was very instructive. As Beyer (1995: 25) suggests, a formative review “requires no magic number of participants” so, the two experts’ feedback, along with the sessions of the six teachers, was enough to detect the most important issues that are related to the checklist usability, practicality and effectiveness at this stage of its development.

5.10.2.2 The Instruments Used and Their Purposes

Unlike one-to-one reviews where participants saw the checklist for the first time at the beginning of the sessions, the small group participants were sent the questions and the checklist in advance because they were required to be familiar with the checklist and its assessment process beforehand. Three instruments were used to collect data: feedback questionnaire, observational log and a presentation at the beginning of the sessions. The

presentation goal before the sessions was to familiarize the participants with the topic of the study, its research questions, materials development and evaluation besides the checklist rationale, sources and use. The purpose of the small group sessions and instruments was to anticipate the possible problems that may obstruct the use of the checklist in its semi-final version during the field testing. The questions were designed to check three main areas in the checklist: effectiveness, usability and appeal (see Appendix J2) for the small group review debriefing questions. The observation log was also used to record the duration of the materials evaluation, the problems facing the participants while using the checklist in the evaluation and the other comments that were raised throughout the review sessions (see Appendices J5 & J6).

The session was conducted on January 24th 2017 at 9:34 am in the College of Technology Foundation Programme offices. The two experts were able to complete the evaluation of the materials being used in the Foundation Programme within a reasonable time: (Expert 1) 40 minutes and (Expert 2) 30 minutes. Two main points were raised while using the checklist: the use of the main category “literature review” in part A of the checklist (Expert 2) and the “rationale of the construction of the checklist items” (Expert 1). Another comment that was shared by both experts was the length of the checklist. The next section will discuss their feedback in details with regard to each question in the debriefing questionnaire.

5.10.2.3 The Experts’ Feedback (Small Group Review)

The experts’ session revealed some of the usability problems in three main areas: the rating scale, the wording of some categories and items and some language problems that resulted from the changes made in the previous version, as well as the researcher’s own focus on the headings and content of the checklist, rather than the linguistic matters. The first three questions under the heading of effectiveness aimed at discovering the usefulness of the checklist, its weaknesses and the solutions for such weaknesses from the users’ point of view (their full answers are available in Appendix J4). In the first question, both experts agreed that the checklist “does cover several aspects of concern” or “most important aspects of evaluation” (Experts 1 and 2 respectively). Regarding the second question, about the checklist weaknesses, both mentioned language errors “few errors in word order” (Expert 2) and “several grammatical and linguistic errors that are made which will also contribute to the reluctance of participants to finish” (Expert 1). The second comment by (Expert 1) was about “numerical scale assigned to measure

differences in responses”. Their suggestion to solve them was through “proof reading” (Expert 1). Clearly, this problem can be solved and avoided when design-based research is conducted through a team of researchers or through spending more time on revisions between each cycle whenever this option is possible and does not clash with meeting the researcher’s important deadlines. To solve this dilemma Baker & Alkin (1973) propose that “there must be support that permits the conduct of parallel treatments over a length of time sufficient to establish differences” but unfortunately “resources of such a magnitude are not often available” (Baker & Alkin 1973: 392). The researcher has to analyze the results of the sessions and write a report and incorporate all the possible changes as soon as these sessions finish.

The experts’ responses to the usability and practicality questions were also helpful and illuminating. In terms of ease of use, both agreed that it is easy to use, but each one of them attributed its clarity to different reasons. Expert 1, for example, thought that it was easy because “there are only three values from which to select” which suggests that keeping the three point rating scale will be useful when using the checklist in teaching materials evaluation. In spite of that, this expert thought that the “overall length” and “the opaque rating scale” may “discourage or hinder the probability of a completed checklist” suggesting that the change will be in the wording of the scale not the number of points. On the other hand, Expert 2 thought that it was easy to use because “it is categorized into clear dimensions”. In reply to question two, Expert 2 presumed that “previous knowledge is necessary” to help in using the checklist, whereas, Expert 1 suggested that at this stage, the checklist still had “superficial flavor in the construction of the checklist items and questions”, which means that greater efforts to modify it had to be exerted by the researcher to work on its main categories and items, before the field testing cycle. Both experts suggested solutions to solve the usability problems. For example, Expert 1 thought that “suitable texts or books should be included with the transmission of the checklist”. Expert 2 also suggested that the researcher had “to give instructions to the evaluators on how to use it” to facilitate its use.

The last four questions were on the appeal of the checklist to its future users (see Appendix J2). In reply to the first question, Expert 1 thought that the checklist still “does not match the complexity that its ideal respondents would most likely find attractive in that it does not accurately reflect all of the different opinions they may hold about the subject matter”. This also suggested more consultation with subject matter experts before

the field testing. He suggested that the checklist “needs to be significantly improved”. In spite of that, this expert praised the “detail of the criteria in each section or component, as well as the selection and total number of possible components or evaluation areas that are listed on each page.” Expert 2 thought that “it is practical” and it “covers the most important issues related” to “materials evaluation” though it is “a little bit long”. On the final question of this section, the suggestions for improvement are listed by both experts. Expert 2, for example, suggested “to shorten it” and Expert 1 stressed the improvement of the “internal or face validity” of the checklist. All of these recommendations were considered through the revisions made to this version of the checklist and which will be expounded next.

5.10.2.4 The Changes Made to the Checklist Prototype 2 (The Experts’ Sessions)

In order to refine this prototype or version of the checklist, most of the proposed changes above were considered. All the language and typing mistakes were corrected, such as the missing verb “do” in the main categories and the typing mistakes of some words. The rating scale was kept as a 3 point scale, but the wording of the scale was changed into 3 options: yes, no, not sure. The items in the quick evaluation checklist in “what to look for” column were organized and numbered to make them clearer for the evaluators. Also all the categories and items were revised to make sure that they were compatible with well-known evaluations conventions. For instance, the three main categories in the literature (second language acquisition principles, teaching principles and curriculum design) were explained in the short summary provided with checklist for next session. The rationale for selecting these three categories were also added. Some of the items that were considered extra were deleted, such as items (g, h & i) in the second language acquisition principles. These items were deleted because they were covered in other sections of the checklist or because they convey similar meanings with other items in the checklist other sections. The phrase “literature review” was replaced by the term “research” as it was more appropriate and comprehensible for the users. Furthermore, the sources and the rationale behind every category and every item in the checklist were acknowledged, as explained in chapter four.

5.10.3 The Small Group Reviews (Teachers)

Following receipt of information regarding the class responsibilities of all the teachers in the English Foundation Programme, the teachers who were teaching different proficiency level classes (intermediate, pre-intermediate, elementary and beginners) were contacted.

Some of them were on their holidays and others were busy. Despite that, six teachers agreed to be part of the small group review sessions. Those who agreed to participate, were sent an email with the checklist and the debriefing questions. Then the time and the place of the sessions were specified in concurrence with the participants' suitable time and schedules.

5.10.3 .1 Procedures, Sampling and Participants

The small group review for teachers' session was conducted in Salalah College of Applied Sciences computer lab. The teachers were invited through emails and reminded several times, but on the date of the session only four teachers were able to come. So two sessions were conducted instead of one. Though these procedures took more time, but the sessions were very useful and informative. The first was on January 25th, 2017 and the second on January 29th, 2017. Each session started with a detailed explanation through a Power Point presentation of the materials definitions, development, evaluation, as well as the use of the developed checklist (see presentation slides in Appendix J1). Four teachers participated in the first session. Teacher 1 is from Syria with 25 years of experience in teaching. His session started 11: 36 am and ended at 12: 30 pm. He used the checklist to evaluate the teaching materials used for (intermediate) students. Teacher 2 is a female from Canada also with 25 years of teaching experience and she evaluated the pre-intermediate materials. Her session started at the same time but finished at 12: 09 p.m. Teacher 3 is from the USA with 6 years' experience and he was teaching two levels (elementary & intermediate). His session also started at 11: 36 and ended at 11: 58 am. Teacher 4 who was teaching and evaluating beginners teaching materials is from Australia with 15 years' experience and his session started at 11: 40 am and ended at 12: 20 pm. The fifth teacher is from Zimbabwe with 22 years of teaching experience and he was also evaluating beginners' level. The session started at 12: 05 pm and ended at 12: 45 pm. Finally, Teacher 6 is from America with 30 years' experience and he was also evaluating beginners teaching materials. The time allocated for teaching materials evaluation using the checklist for the six teachers was: 54 minutes, 33 minutes, 22 minutes, 40 minutes, 40 minutes and 25 minutes with the average of 35.7 minutes.

5.10.3 .2 The instruments used in data collection

The teachers' sessions followed the same procedures as the experts' sessions of the small group review cycles. The researcher presented a presentation at the beginning to explain the development and the use of the evaluation checklist, followed by the teachers'

evaluation of the teaching materials. An observational log was used to record the comments and the problems raised by teachers while using the checklist (Appendix 5). A second observer helped in the first session, but in the second, no help was needed as there were only two teachers in the session and it was easy to write their comments on the checklist. After finishing the sessions, the teachers answered the feedback questionnaire that consisted of three headings (effectiveness, practicality and appeal) and 10 questions (see Appendix J3 for complete summary of the participants' feedback). The answers to these questions, along with teachers' annotated comments on the checklist itself and the researcher's observations are discussed next.

5.10.3 .3 The Teachers' Feedback From Small Group Reviews

The teachers were given the session packages and were asked to select the materials they have taught in the previous semester. The researcher had already prepared copies of teaching materials from different proficiency levels in advance. So, each teacher had the teaching materials, the checklist and the feedback questionnaire. The first question raised by one of the teachers and recorded in the observation log was: "is the checklist suitable for all levels?" The researcher's answer was yes and their ability to use it to evaluate the different proficiency levels provided the practical answer for this question. The other problem that faced the teachers was the scoring criteria. Also, Teacher 2 asked for the simplification of "the terms used", which also concurs with one of the experts' comments in the previous expert session of the formative review. This led the researcher to revise all the categories and items in the checklist to simplify the difficult and vague items. For instance, the abbreviation L2 was changed into second language. The terms 'input' and 'output' were also changed in the main headings (see the modified prototype in Appendix L3). Also, the answer for the question "is the checklist for the whole package or just the main textbook or coursebook?" that was raised by teacher 5 at the beginning of the session was 'yes'. In fact, the quick evaluation part of the checklist is about the teaching materials as a package rather than a single textbook. Some teachers wrote their notes and comments in the last column of the checklist entitled "the evaluators' notes". Some other comments that were not related to the checklist were about some of the problems the teachers face like the higher level of the textbooks for "zero beginners" and the mismatch between the reading and writing texts and the final exams. This particular point is debated many times in the Colleges of Applied Sciences, but nothing has been done to solve it. Commercial textbooks will have to consider such issues and suggest the type of the assessment that

will help to test not only the achievement of the students, but also reveal their proficiency levels by the end of the course.

The teachers' answers to the first question in the feedback questionnaire about how the checklist helps in materials evaluation was very interesting (see Appendix J3 for the teachers' answers). Teacher 6 saw it as "a guideline to make changes and recommendations for future curriculum changes" and Teacher 5 thought that "it highlights critical issues related to T.B. evaluation". Teacher 3 thought that "it helps" him to "understand the needs of an ESL textbook and have a standard by which to judge them" so it is considered as a needs analysis instrument as well as an evaluation tool. Teacher 4 thought that "it helps analyze the components needed to make course material become functional" which advocates the importance of regular evaluation of the teaching materials in the English Foundation Programmes that is rarely conducted in the Colleges of Applied Sciences. Teacher 2 reply to the same question was different in the sense that it suggested a solution to make the checklist more practical so that it can "be more easy to recognize at a glance" [sic]. This advocated the importance of the checklist layout as well as its content to simplify the checklist for its users, especially teachers. In reply to question two in the effectiveness section and which asked about the checklist weaknesses, the researcher got six different views. Teacher 3's and 4's comments were about the rating scale and how to improve it: Teacher (3) suggested to start with a phrase "on a scale from 1-3 how much do you agree"? Teacher (1) complained about "too many things put under one item" whereas Teacher 2's focus was on the checklist's visual appeal and suggested that "sub-headings on checklist could/should be more block by block clear from first glance". Teacher 5's and 6's comments were more on content and the general aim of the checklist compared to other teachers who focused on the layout and meanings of categories and items of the checklist. To solve the problems identified by the teachers, question three purpose was to get as many solutions as possible. Teacher (1) recommended "designing a more elaborate one" whereas as Teacher 2 still adhered to the visual layout where she recommended to divide categories and items into clear "sections" like "contents of books" and "classroom interactions". This comment calls for more salient headings and sub-headings, which was considered in the checklist adjustments. But, as the checklist development is based on the conceptual framework, such feedback was dealt with carefully in order to keep the main categories as they are demonstrated in the framework. The main purpose of the study is to find the appropriate method for

teaching materials evaluation checklists development rather than reproducing structures and contents of previous developers and researchers' checklists.

The usability and practicality section also had three questions. Generally speaking, all the teachers managed to finish their evaluations in spite of their complaints about a few aspects of the checklist. As a result, the teachers' answers to question one were positive. Apart from Teacher 2 who thought that previous knowledge and clear instructions were required to facilitate the checklist use, the rest of the teachers thought that it was easy to use and would require only "experience as a teacher" and "longer teaching experience" (Teachers 3 and 4 respectively). More suggestions appeared in the answers for question 3 which ranged from using "simple and straightforward ideas" as suggested by Teacher 1 to just familiarizing "with the materials" which the teacher or evaluator is "about to evaluate" according to Teacher 6 and Teacher 5. It was also suggested by Teacher (5) to have "formal TEFL training" as well as "brief instructions on how to fill out" the checklist. The checklist and "its purpose" was suggested to be explained before the evaluation by Teacher 3. Teacher 2 proposed "a verbal and visual explanation of the checklist parts and terms". Based on the above recommendations, it was decided to provide a short summary as an introduction to the checklist and to simplify the terms, especially the ones in the Second Language Acquisition principles. This suggestion was later provided in form of a summary of the rationale and the use of the checklist for field testing. The last section of the feedback questionnaire on the checklist's appeal to the potential users consisted of four questions. The first was if the users find it "interesting and attractive" and all the teachers found it interesting, but Teacher 4 questioned its attractiveness. The issues that were interesting to them were its "close reference to key issues regarding textbook evaluation" (Teacher 5). Teachers (1, 2, 3 and 4) thought that it was "extensive, clear and well organized" whereas Teacher (4) liked its "linear layout". Teacher 6 was interested in question three in the quick evaluation checklist as it is related to students, teachers and materials. Teachers (2 & 3) expressed some of the issues that may affect the appeal of the checklist such as "small print", as well as "lack of not addressing the impact that language carrier all the cultural aspects of L2". The teachers' answers to the last question exemplified important suggestions that would be considered in the third prototype of the checklist, as suggested by the role of formative review, which may "identify redundancies or areas of omission in the development process" (Baker & Alkin 1973: 413). Teacher (1), for example, recommended to "use as less items as

possible” and as result, some items in the needs analysis categories were removed and merged. Removed items were mostly from institutional needs which included items (1) and (4) in section “a” and item 3 in section “b” as they were repeated in other sections of the checklist. The first two items (a and b) in second language acquisition principle were merged. Teacher (2) list included: “large print, secondary sheet for longer responses, clear, functional and useful sub-headings and much shorter topic sentences for each question/ area”. She also thought that the “overall organization must be well thought out”. Teacher (6) also suggested that “it is important to address the language-culture dichotomy” because “language cannot be separated from culture”. As a result and in response to these comments and feedback, the checklist was amended for the third time to make it more practical and visually appealing. The version of the prototype resulting from these small group review sessions was used in the final stage of the teaching materials evaluation checklist testing.

5.11 Field Testing of the Checklist Prototype 3

5.11.1 Introduction

Through the field test stage, the developers would have three main goals: to “confirm the revisions made in previous formative” reviews; to “generate final revision suggestions” and to “investigate the effectiveness” (Tessemer, 1993: 137). In this cycle, which is more complicated than the previous cycles, the six Colleges of Applied Sciences were the test setting. First, the emails of the participants were obtained from the Foundation Programme coordinator. Then, the participants were contacted through emails and phone calls. The participants were contacted individually and in groups so that they were introduced to the study and all their questions could be answered immediately. The chatting groups were created to facilitate communication with them and in order to make sure that the data were collected through a reliable and valid procedure. The main problem in this cycle, was the fact that the six colleges are geographically far from each other and any attempt for the researcher to be present physically during the field testing of the checklist use was very difficult and time consuming. An online field test was envisaged as a solution, but this option also encountered some problems. The first problem was being unable to use Skype as it is not allowed in the country. The second problem was finding an alternative that is easy to use by participants and can be used in the Omani setting. The only applications that are free for the participants are Hangout and

IMO: the latter was chosen as it is accessible by all participants in the five Colleges of Applied Sciences.

5.11.2 The Field Testing Sessions' Procedures and Participants

When agreeing on the date and time of the sessions the participants were sent the following documents: the evaluation checklist, instructions on how to use it, a summary about the study and the feedback questionnaire (see Appendix K1). For this cycle of data collection, ten participants from the six colleges and Dhofar University were contacted and the focus was on coordinators, as their position requires their involvement in materials' selection and evaluation. Two experts from Dhofar University were also contacted, but only one was able to conduct the session. The other expert was busy during the time of data collection and later apologized for not being able to participate in the session. This cycle was challenged by many obstacles, as mentioned earlier, due to the geographical locations of these colleges. They are scattered in different regional areas where travelling may take a longer time than that allocated for this cycle. As a result, five of these sessions were conducted online. Some participants agreed to use IMO, and even with IMO, a few of them faced some problems due to slow connectivity networks in their colleges. Also, two female coordinators, expressed their discomfort with online sessions, but they recommended substitutes who were willing to do the sessions online and they provided their contact details. In spite of these obstacles, the researcher was determined to finish all the sessions in order to test the checklist in the six colleges.

Three of the 10 participants were coordinators who were involved in the teaching of the materials besides their administrative role as colleges' representatives in the Ministry of Higher Education. The rest were teachers who were also teaching courses or who had taught general English courses in the English Foundation Programme. Seven of them were males and two were females from different backgrounds and with diverse teaching experiences that ranged between 4 years and 24 years. All the participants were from the Colleges of Applied Sciences apart from one expert from Dhofar University. The participants were also from different countries and though most of them were from Oman, others were from the United Kingdom, Sudan and India.

During online sessions, the participants had the teaching materials that they had taught or the ones that they were teaching currently. They were sent all the documents in advance with a short summary of the study and instructions on how to use the checklist and the

procedures for the data collection sessions. The instruments used throughout these sessions were based on Tessmer's (1993) model and guidelines on how to conduct formative reviews. The feedback questions were basically about the problems that occurred when approaching or using the checklist, while the observation log focus was on the participants' performance, behavior and the time spent to finish the teaching materials evaluation using the checklist. There was also a section for the researcher's notes for every session.

5.11.3 Feedback Questionnaire (Field Testing)

As this is the final review in this study of the teaching materials evaluation checklist, three instruments were used to collect data during the review sessions. The feedback questionnaire had six questions that investigated the problems faced by the users, and how they were dealt with, the required training for the checklist use and if it was appropriate, for their institutions' needs, to evaluate the English language teaching materials. Thus, the results of this cycle were obtained first through the participants' answers to the feedback questionnaire, second through the observations recorded using the observational log and finally the researcher's notes on each session. Results and discussions of the participants' answers for each question in the feedback questionnaire are represented next.

Q1. Did you have problems approaching or start using the checklist?

Most of the answers for this question were positive responses apart from two instances that were related to technical problems with the checklist design such as Participant 2 who "couldn't read the second part of the checklist (detailed checklist) when the file opened" as he "had to zoom" it and Participant 6 who had weak internet connection. Participant 6 also had to ask about few headings and items in the checklist "few bullet points in the checklist" as he said. Likewise, Participant 8 said that he "did need some help understanding some of the questions, mainly to give a more accurate answer". So generally speaking, the users did not have major problems when they started using the checklist.

Q2. Did you have problems while using the checklist?

The second question goal was to discover while-use problems. Some of the usability problems were simple like "changing between pages" (Participant 1), "some columns consistency" (Participant 3) and "some typos need to be checked and corrected

throughout” (Participant 5). On the other hand, some other problems, required adjustments in the checklist itself. For example, Participants 8 and 9 asked for “more options” for the rating scale.

Q3. What kind of problems did you face?

Some of the problems mentioned in question three can be classified into two groups: the first group is related to the content of the checklist and the second is associated with its use. For example Participant 2, wrote that “some questions were unclear and he provided this example: “follow the description from listening texts” in (teachers’ needs item b No. 4). He also suggested that “some Qs impossible to answer as they contained multiple Qs” such as “accents and real conversations”. Participant 2 referred to “tips for speaking and writing” as a problem because the evaluator has to evaluate two items at the same time instead of one. There was a comment from Participant 4 that teachers’ needs “and its statements/phrases” has a major problem. He thought “that most of them were not related to teacher’s needs (preferences, beliefs, personality/ identity etc.)”. He also added that “combining two different terms (teaching and learning) in one statement was confusing.” Another participant (Participant 9) “felt that there are so many questions” and he thought that teachers would not be able to spend “all that amount of time in answering them unless a specific session is conducted and prepared only for this purpose”.

In terms of the checklist use, Participant 7 thought that “some statements were not shown as complete statements on the Excel sheet”, which meant that the checklist had to be edited in its final version to avoid such design errors. Participants 8 and 9 complained about the checklist rating scale as their comments suggested. For example, Participant 8 wrote “I would’ve preferred a Likert-scale continuum to answer some of the questions”. The same opinion was expressed by Participant 9 saying that “there is no scale to add the exact opinion”. This issue was dealt with based on the whole formative review cycles as it occurred in almost all of them. So, the overall feedback was used to improve the checklist rating scale. The only expert in this field test (Participant 5) referred to “some typos” that “need to be checked and corrected throughout” as well as “consistency...of some terms (e.g., L2 for English language)”. He also added that “participants should be told at the beginning of the session about the type of evaluation (i.e. pre-and/or post).

Q4. Is additional training needed on how to use the checklist?

Regarding the fourth question and apart from two participants, most of them thought that conducting the teaching materials evaluation using the checklist would not require any additional training. Of the two who did think the use of the checklist required training, Participant 4 said that “if the checklist items are modified with more specific items, there will be no need for extra training.” Participant 9 also thought that “there is a need for that because not all teachers are really aware of curriculum design and also, teachers need to be informed about the importance of this checklist in order to answer it.” So out of the ten participants, only two thought that additional training is recommended. The rest said there was no need for additional training and thought that its use would not be difficult to other users (see Appendix K3) for full answers. Based on these views, it seems that what the teachers and users need more in the teaching materials evaluation checklists is clear content, instructions and guidelines of their design and use.

Q5. Are more guidelines needed for the use of the checklist?

Four participants’ answers to the fifth question were in favour of providing general guidelines for the checklist use. Participant 1 suggested “a brief outline to the contents/ topics asked”, while Participant 4 thought that “the researcher needs to clearly define some pedagogical terms such as course, syllabus, and curriculum”. Participant 8 recommended providing certain guidelines “in some cases” for those who may need them and he thought that they should be supplied “only at the start” of the evaluation and not throughout. Also, Participant 9 thought that “there should be a clear guidance and instructions informing the teachers to choose the target book”. The rest of participants (6 teachers) said that there is no need for additional guidelines on the use of the checklist.

Q6. Does the checklist satisfy the users’ need (in your institution) for evaluating teaching materials? How?

All the participants thought that the checklist will be useful in evaluating and selecting teaching materials in the English Foundation Programmes. Their explanations on how it might help were different. Participant 1 thought it is a “helpful” tool for evaluation, whereas Participant 2 thought it will be useful for teachers when they “need to evaluate if material facilitates the move from general to academic English”. Participant 3 thought that the checklist has “Clear options. Simple statements. Understandable language and instructions”. Participants (4, 8 & 9) agreed that it is useful to some extent and proposed

some recommendations for the checklist's efficiency. Participant 4 thought it will need "more modification (especially the section of Teacher's Needs). He also thought that "the checklist items need to be organized according to their categories (layout and design, curriculum, language skills, language content, topic content, activities, methodology, learner's needs, teacher's needs, institutional needs)." Participant 8 stated that "using this checklist can help us narrow down our choices when choosing the best teaching materials and textbooks" but he thought also that the checklist is "a bit long and some questions might be more related to teaching materials (supplementary, rather than textbooks)". The suggestion by participant 9 was "adding couple of questions targeting the ability of teachers to use the technology requirements- in the text book- to activate the lesson." Participants 6 and 7 thought that the checklist provided "valid points to consider when choosing a textbook for the Foundation Programme or any other ELT programme" as well as including "the main points to be considered when evaluating any materials." Finally, the expert in this cycle, (Participant 5) thought that it can be useful "to a very good extent" and that "future research could move further and include other levels of analysis (e.g., consumption and production)" as the "current study only focused on the course book content level of analysis."

5.11.4 The Observational Log (Time, Performance and Behaviour)

Time

The time taken to finish the evaluation of teaching materials using the checklist ranged between 24 minutes and 1 hour and 24 minutes with the average of 46 minutes for all nine participants. Participant 10 was not included, because she did only the quick evaluation checklist as the total score was less than 80% in the quick evaluation. This was a satisfactory result considering that some evaluations may take weeks and months, a time that is not available for practitioners in the English Foundation Programmes. When comparing the economical aspect of the checklist and its comprehensiveness, it is hoped that the checklist will help in reflecting the real value of the teaching materials for both teachers and coordinators within a short period of time and with satisfactory evaluation results.

Participants' Performance

Generally speaking, the participants' performance was acceptable. The main focus for the researcher during the observation sessions was to spot any problems that may impede the use of the checklist. So, the performance indicator here is the ease or the difficulty of teaching materials evaluation checklist use. The first session was with two participants (1 & 2). This session went without problems apart from two instances mentioned by the participants in the feedback questionnaire. These two problems were: finding difficulty in approaching the detailed checklist due to a technical problem for the first participant and the font size for the second. Participant 7 encountered the same problem about the font size of the detailed checklist and the participant was directed to the place where he can enlarge the view of the checklist. Both Participants 3 and 4 had the same comments about the quick checklist, which included a few spelling and layout issues as well as some missing words due to the new changes made in the checklist after the small group review sessions (e.g. category 6, item 1). Also they said that they did not know the price of the teaching materials package. Participant 6 also referred to these simple typing errors which may make teaching materials evaluation difficult for some users and evaluators. Participant 9 session went smoothly and did not encounter any problems.

As for the detailed checklist, Participant 3 had a few comments. First, the font size of the second sheet or the detailed checklist was small. He also commented on the use of authenticity in teaching materials and suggested to devote a section on "what is authentic". Other technical issues in second language acquisition principles included: items "b" options font size were very small. The same comments were given about other categories such as "Curriculum Design" where some focus was drawn on the items with errors such as "b. items 2 (in the teachers' book instead of textbook). Few corrections were also mentioned under the needs analysis section to include: Students' needs: section (b) item 1 and section (a) items No. 3 and 4 spelling of (themes) and (students). In the teachers' and institutional needs, the verb 'do' was missing in some evaluation questions. Participant 4 mentioned also some of the problems like capitalizing "year of publication" as well as Ministry of Higher Education. He said that font is not consistent and suggested that 12 would be good for all. He also thought that items 3 and 5 in the teachers' needs section were misplaced, as item 3 is compatible with assessment (under the course, not teachers' needs) as well as item "current and up to date". In fact these items were based on brainwriting sessions with teachers who thought that they would prefer up-to-date

materials. This participant also had problems on deciding about the number of themes or activities in the materials. For example, research skills, where only two units have research-based activities and if this is considered enough coverage as well as the availability of satisfaction surveys for students but not for teachers. Participant 6 inquired about what an 'appropriate size' meant in the quick checklist. She also referred to some items that need clarification, as in section 1 (SLA principles... "f" item 2: share does it mean to talk about?), and section 2 (Teaching principles... "a" item 1: meaning of "throughout the materials" ...) suggesting to use "in almost all units" instead. She also asked, in "teachers' needs", about principle "a" item 1, "current and up to date", and what did it mean? In the second language acquisition principles item (f) and in teachers' needs section, she asked about personalized activities and if it means to "talk about themselves". She also mentioned the importance of the evaluator's notes column for the evaluators and for the other users and stakeholders to include any concerns that resulted from the checklist use or the materials' evaluation.

Participant 7 mentioned some missing items due to the small height of some rows (e.g. SLA principles "a" items 1 and 2). The participant could solve these problems by himself when the cell was activated. In general, this session was very easy for that participant apart from a technical problem due to coding error in the checklist design where there were some answers which were already in the checklist, so the participant had to select the appropriate answer again for every question. The problem was explained to the participant and he was told that it was not intentional but this problem occurred because of the Excel sheet design. Participant 8 commented that the quick checklist was focused on different items especially the price item. That item seemed to raise some concerns for the teachers as they do not know how much the materials cost. This participant searched for the price on the internet in advance of the session and estimated that it was about 120 OR per student and he thought that this was very expensive. In spite of that, this item was kept as it is a very important part in the evaluation and selection of the teaching materials for the English Foundation Programmes. It would also encourage involving the teachers and make them appreciate the value of the materials pedagogically and financially. Participant 10 session was quiet but with a slower process due to the slow computer. The session stopped also at the quick checklist as the score was less than 80%, so the participant did not continue the close evaluation section of the checklist. Also, Participant 5 Excel sheet, was not compatible with the version used to design the checklist and did

not open, so the participant had to use the researcher's laptop to complete the evaluation session. The drop-down list of the three options (yes/no/ not sure) also did not work in the laptop, so he had to use the touch screen instead of the keyboard. Also, Participant 5's evaluation session was affected by the font size of some options from the drop-down list as it was very small. The same participant thought that the word "level" needed to be specified by adding 'language proficiency'.

It was noticed also that the participant's previous experience with teaching the same materials presented in the session affected their evaluation and the selection of items in the checklist. There was a request for more details in items that included "teachers' and the learners' role" and to explain more what is meant by teacher or students role. Other notes were about "teachers' needs" principle b: item 2 and 4 and the meaning of "follow descriptions" as "it needs to be clarified". There were also some typos, an incomplete sentence and the need for consistency in the use of terms (e.g. second language and English language). A final note was about the very important issue of context as participant 5 questioned the difference between learners' context and teachers' context. He thought that teachers' context was about perceptions of teachers, cultural background and language teaching background (novice vs. experienced). So the question would be more appropriate if it were rephrased as "to what extent are materials suitable to the learning context/ teaching context instead of teachers' or learners' context.

Participants' Behaviour

The aim of this part was to observe participants' engagement and acceptance of the teaching materials evaluation checklist in general. So, the goal here was not to get a detailed measurement of the users' experience while using the checklist, but rather to make sure that the developed checklist was mostly understandable, interesting and functional for its potential users. The usability of a product is defined by the ISO (International Organization for Standardization) to mean "the extent to which the product can be used by specific users to achieve specific goals with effectiveness, efficiency, and satisfaction in a specified context of use" (Scholtz 2004: 1). The users' experience which "goes beyond normal usability and functionality aspects of products by incorporating user's feelings and emotions towards these products" (Allam & Dahlan 2013: 29) would require another phase of design-based research and is not the focus of this cycle of

assessment in this study. The observation of the participants' behavior while using the checklist differed from one participant to another. Some participants were very quiet and they could finish the evaluation in a relatively short time (24 minutes), others, who asked for explanations for every difficulty they encountered, took more than an hour. Some participants wanted to read every item with the researcher, but unfortunately, the aim of the session was to focus on the use of the checklist rather than discussing it. So the participants were asked to continue their evaluation quietly unless they faced a major problem as they could eventually write all their comments and the problems they encountered in the feedback questionnaire. For example, Participants 1, 2, 7 & 9 were very quiet during the evaluation sessions, which enabled them to finish in a short period of time compared to other participants (37 minutes, 41minutes, 24 minutes & also 24 minutes) respectively. On the other hand, some participants wanted to do the evaluation as pair work activity where they could discuss their thoughts with the researcher and ask for clarifications. Participant 5, for example, spent an hour discussing the quick evaluation checklist prior to the session as he thought that more items should be incorporated; it was later understood by the researcher that the detailed invisible checklist in the Excel sheet may confuse the evaluators' who may have thought that there is only that short checklist, which does not cover all the items for conducting a trustworthy materials' evaluation. As a result, consideration would be given to making both checklist parts (the quick and the detailed) visible for users so they can see both checklists. Participants 3 and 4 had had some discussions while using the checklist, but later Participant 3 decided to write his comments in the evaluators' notes column and Participant 4 continued the evaluation after referring to some items in the checklist. Participant 8 tried to look at the materials in front of him checking them against the items in the checklist. Finally, Participant 6, needed to discuss the checklists items loudly to check the meaning of some items, suggesting that using the checklist among a group of teachers would be a useful method for evaluating and selecting teaching materials. Participant 10 checked the availability of some items in the materials during the use of the quick evaluation checklist, but she could not do the detailed checklist as the score in the quick checklist was only 75% which did not allow proceeding to the detailed evaluation checklist.

The above data collection instruments helped to discover most of the problems as well as the strengths about the teaching materials evaluation checklist. The usual processes for

field testing in this study as proposed by Tessmer (1993: 143) included the following actions:

- Observe the intervention
- Administer the survey
- Debrief participants to detect any problems or the need for training
- Review the data
- Repeat and/or report

As can be noticed from the above activities, even field tests can be repeated when needed. The instances where the product or intervention field testing can be repeated usually occur when “a significant number” of participants cannot use the instrument, when they use it “incorrectly”, when they cannot use the instrument “without some form of extra help” or if it is “boring or unchallenging” (Tessmer 1993: 147). Fortunately, the developed checklist did not need to be repeated as all participants were able to use it. The previous formative review stages helped to avoid such issues in the field testing. Despite that, the evaluation checklist as an incessant project for materials selection and evaluation and can be constantly improved and updated using any of the formative review techniques, especially field testing. The final changes made on the teaching materials evaluation prototype checklist are clarified next.

5.12 The Changes to the Teaching Materials Evaluation Checklist Prototype 4 (after field testing)

The changes made to the checklist prototype after field testing were fewer than other sessions as most of its major problems were identified in the previous reviews. The changes made in this prototype were related to reducing the number of items to make the checklist more practical as well as adjusting some sentences in the teachers’ needs heading. Some of the items were also modified to make them simpler. A summary of the checklist use and instructions would, at a subsequent stage, be provided with a brief outline about the content. Also, some terms were explained such as ‘course’, ‘syllabuses’, and ‘curriculum’. The general layout and the font size were also improved. The rating scale was kept as it is because the three options were considered better than five options in term of reducing the confusion and time spent on doing the evaluation. Some suggested only two choices: ‘yes & no’, but adding the third option: ‘not sure’ is necessary especially when one of the items is not covered in any of the units of the textbooks. In

such situations, the evaluator can select that option, as happened in one of the sessions in the field testing. Also, using only yes/no options would turn the evaluation procedures in the checklist into a content analysis rather than an evaluation process. As a result, the three options were retained in the final prototype of the checklist.

Chapter 6 Teaching materials Evaluation and Design-Based Research: Results and Reflections

6.1 Introduction

As explained in chapter 1, institutions such as Colleges of Applied Sciences in Oman are an example of how an important issue like teaching materials evaluation can be neglected when it comes to purchasing and using English language textbooks. The problem of some educational evaluations lies in the fact that they are dealt with as whole packages and as part of accreditation audits or general evaluation of a whole programme. As Baker & Alkin (1973: 403) state, the “conceptions of evaluation have for the most part been addressed to the evaluation of total educational programs with specific context designation rather than the evaluation of instructional products being devised by developers.” Despite the importance of such quality procedures (quality assurance inspections), most of the time, the top-down view may not help to inform the appropriate changes and adjustments these programmes need. In such situation, using an evaluation instrument or even one of the formative review techniques could help to save money and time. George & Cowan (1999: 32-33) explain the purposes of evaluation and its importance, indicating that the findings of an evaluation process can be used for many intentions which include reinforcing “a need for a change... informing review and debate, discovering unperceived needs, establishing an unperceived need” and changing “attitudes”. The idea of evaluating the materials or the course before use or after use can be considered “the greatest service evaluation can perform” as it identifies the “aspects of the course where revision is desirable” (Cronbach, 1963 cited in Griffee & Gorsuch 2016: 6).

The evaluation of teaching materials needs to be done using both inductive and deductive methods. By doing that, evaluators can ensure that they have focused on the teaching materials as a product and as a process besides involving most stakeholders (teachers, learners and authorities). It is true that “materials evaluation is initially a time-consuming and difficult undertaking”, despite that when conducting materials evaluation, it can “help to make and record vital discoveries about the materials being evaluated” besides helping “the evaluators to learn a lot about materials, about learning and teaching and about themselves” (Tomlinson, 2013a: 44). Also, Balachandran (2014: 85) states that materials evaluation can “motivate the teachers to reflect on the gaps in the existing textbooks and this would pave the way for incorporating newer strands into the course package when a

revision is undertaken.” The accessibility to such evaluation tools facilitates the accomplishment of these purposes and many others that are mentioned in the next sections. Teaching materials with their multiple roles, are very important in the English language programmes. So their constant improvement requires a clear plan and a practical evaluation instrument which becomes a must for these programmes’ development. Richards (2001: 1) explains the role and importance of teaching materials in the English language programmes (discussed in thoroughly chapter 2):

Textbooks are a key component in most language programs. In some situations they serve as the basis for much of the language input learners receive and the language practice that occurs in the classroom. They may provide the basis for the content of the lessons, the balance of skills taught and the kinds of language practice the students take part in. In other situations, the textbook may serve primarily to supplement the teacher’s instruction. For learners, the textbook may provide the major source of contact they have with the language apart from input provided by the teacher. In the case of inexperienced teachers textbooks may also serve as a form of teacher training they provide ideas on how to plan and teach lessons as well as formats that teachers can use.

Skierso (1991 cited in Bülent 2006: 22) postulate that “textbooks evoke a variety of emotions in their users” and “no teacher is entirely satisfied with the text used, yet very few manage to teach without one.” Despite the importance of teaching materials, the authors of commercial textbooks and famous publishers do not conduct enough research on the needs of various stakeholders. As McGrath (2013: 30) point out, “publishers of ELT materials make most of their money from coursebooks, dictionaries, and grammar books” but unfortunately they “tend to play safe by commissioning new editions of popular series”, which means that stakeholders' needs are usually ignored when developing teaching materials. The materials writers depend heavily on the research findings where they consider only some of the educational research outcomes or ‘trends’ among academics. For example, “most commercial textbooks and resource books display the influence of discourse studies. A representative example is Headway (Soars and Soars, 2005)” (Tomlinson 2003: 133) the core textbook in the Colleges of Applied Sciences in Oman. In another instance, Tomlinson (2014) refers to the same problem, clearly indicating that “publishers dare not risk losing vast sums of money on a radically different type of textbook, they opt for safe, middle-of- the-road, global coursebooks which clone the features of best-selling coursebooks such as Headway and they cut down on non-profit-making supplementary materials” (Tomlinson2014: 7). As a result, most

publishers are “inevitably driven by perceived market needs” not “by syllabus needs” or “learners needs” Tomlinson (2013: 8).

In such a complicated situation, the gap between materials development and materials evaluation creates more problems for the materials users or between the textbooks production and consumption in Harwood’s (2014) terms. Thus, the availability of the evaluation criteria are not only used to appraise the quality of the teaching materials but are also an essential part of their refinement and development. Reaching such targets may help also to force publishers and writers to state the principles, approaches and theories behind the development of their textbooks, rather than vague claims such as ‘multi-syllabus’, ‘communicative’ or ‘authentic materials’. Thus, the teaching materials checklist here becomes more than an evaluation instrument, it is an indicator of the important content and items that are supposed to be included in an effective-materials or a see-through tool of the core elements and components in any teaching material developed for the English language programmes.

For specific evaluations of teaching materials in the English Foundation Programmes, the instruments used have to be both practical and effective because, as Tomlinson (2013: 31) emphasized, making use of checklists as “criteria has become popular in materials evaluations”; Tomlinson gives examples of several checklists that have been used in these evaluations. McGrath, (2013) also sees checklists as the most practical evaluation tools, suggesting impressionistic evaluation first and then close evaluation. In this study, the checklist has been found to be viable for both quick and detailed evaluations of teaching materials, which was the reason for combining both aspects in one comprehensive checklist. This final chapter focuses on reflections on the findings of this study, thoughts about DBR methodology and its contributions to educational research as well as its limitations. The chapter also summarizes the implications of the study as a whole and its limitations.

6.2 The General Processes of this Study within DBR Phases

The design of the checklist prototype went through several phases and stages starting from specifying its sources and ending with its assessment by potential users. For example, the analysis and exploration phase (phase 1) included two major cycles, literature review and investigation of setting needs. Through this phase eight different activities were used, which comprises (i) literature review, (ii) informal discussions with

coordinators from the six Colleges of Applied Sciences, (iii) brainwriting data collection sessions, (iv) brainwriting data analysis, (v) peer review of the analyzed data from brainwriting, (vi) a short survey sent to a representative of colleges' authorities, (vii) the development of the conceptual framework based on literature reviews and setting needs analysis and finally (viii) expert appraisal of the conceptual framework for the checklist sources through five open-ended survey questions. The result of these activities was the preliminary prototypes of the teaching material evaluation checklist. The second phase was the design and construction phase, which encompassed the creation of the first complete checklist prototype. Expert review was used to check the content and construct validity of the checklist. This phase helped to transfer the abstract ideas and thoughts obtained from the theoretical knowledge and the setting needs analysis into a feasible prototype that was used in the following phase to get the required feedback from the participants and users (all these activities are discussed in detail in chapters 1, 2, 4 & 5).

The checklist developed prototype was tested in the evaluation and reflection phase. In this study, a prototype refers to the "preliminary version of the whole or a part of an intervention before full commitment is made to construct and implement the final product" (Nieveen 2007: 90). Also, the type of the prototype that used here is the one that is "continually refined" based on data collection feedback and that evolves "towards a final deliverable" version of the intervention. This type of prototype according to Nieveen (2007) is called 'evolutionary prototyping' to differentiate it from paper prototypes or throw away prototypes that are used in the early stages of the design and discarded when the feedback is obtained allowing for the design of a new version. During the design and construction phase, major activities were used to construct the checklist prototypes, which included designing the initial two prototypes of the checklist based on the conceptual framework first and later combining the two prototypes of 'research and 'setting needs' into one organized prototype. The checklist prototype with its main categories and sub-categories was established. Finally this prototype went through the developer screening using four previous studies on designing materials evaluation checklists in particular and evaluation checklists in general. Formative review with its instruments and techniques was used in data collection and in reporting the participants' feedback, which was in turn used to revise the teaching materials evaluation checklist prototypes. This was the beginning of the third phase (evaluation and reflection) which included four cycles where four techniques of formative review were used (one-to one evaluation, experts review,

small group review and field testing). The synergy between the design and construction phase and the evaluation and reflection phase helped in both refining the design and fixing the problems that had occurred during the checklist development and use. To ensure the success of the formative review processes, clear guidelines were followed, which were provided by two key references; Tessemer (1993) and Beyer (1995). Tessmer’s (1993) provided recommendations for each cycle leading to designing robust data collection instruments as well as data analysis. For example, the goal of developer screening was to discover obvious errors, the expert review and the one –to-one review aims were to check the content and construct validity of the teaching materials checklist. The small group review target was to assess the checklist effectiveness whereas the field testing was aimed at checking the implementability and usability of the checklist. The field testing helped to identify “the constraints of daily practice that impacted the tool’s use and effectiveness” (Richey & Klein 2007: 57) through the various reviews conducted by experts, teachers and coordinators. Also, Beyer (1995: 55) advised that formative review researchers should “collect data at several points in the development process, use multiple data-collection methods, and modify or revise the product at each point in the development process.” He stressed that “even when the changes compromise or challenge our pet theories or ‘ideal world’ of the researchers, they should not hesitate to make the appropriate changes provided by the users. Finally he warned of confusing or equating formative reviews with “editing” as formative review “deals with the substantive more than with the cosmetic” like spelling, colours and paragraphing. Table (8) summarizes the purposes and the methods of formative review method.

The main purpose	Formative review cycle	methods	Instruments used
<ul style="list-style-type: none"> ◦ Validity - Relevance (content validity) state of art knowledge - Consistency (construct validity) format and logical connection 	Cycle 1	Developer screening	Literature review and previous studies in checklists development
	Cycle 1	Expert review	3 previous studies in checklists development Four open-ended survey

◦ Clarity and Appeal	Cycles 1, 2, and 3	Expert review One-to-one review Small group review	The expert survey Data recording log Debriefing questions
◦ Practicality (usable in setting...fit with setting and users)	Cycle 3 and 4	Small group review	Data recording log Debriefing questions
◦ Effectiveness (intervention achieves desired outcomes)	Cycle 4	Field testing	The checklist prototype Observational log Feedback questionnaire

Table (8) The study Formative review purposes, methods and instruments

The use of design-based methodology with its core phases enabled the researcher to answer all the research questions raised within this thesis. The previous processes and phases along with the main research questions are exhibited in Table (9) below.

Analysis & Exploration Phase	Design & Construction Phase	Reflection & Evaluation phase
<ul style="list-style-type: none"> ◦ Literature review(s) to define sources of the evaluation checklist and suggested research areas: English language programmes: English for general purposes/ academic/ business...; learning theories and principles; teaching theories and principles; ELT Curriculum: syllabuses philosophies and models, instructional design; Evaluation: evaluation theories and models, teaching materials evaluation methods; 	<ul style="list-style-type: none"> ◦ Based on the conceptual framework and its main categories and sub-categories, the design and construction phase of the evaluation instrument starts considering the background information (in evaluation checklists in general and teaching materials checklists in particular) from phase one, ◦ The development of the first prototype(s) is 	<ul style="list-style-type: none"> ◦ With the creation of the evolutionary prototype, the third phase (evaluation and reflection phase) was initiated using formative review assessment cycles to validate the developed checklist ◦ Four prototypes of the same teaching materials evaluation checklist were tested. Prototype 1: one-to-one and expert review...Prototype 2: Small group

teaching materials evaluation checklists. ◦ Setting needs analysis: main stakeholders needs (learners, teachers and authorities or institutions) ◦ Developing a conceptual framework (that connects all the researched areas in this phase). Defining conceptual and theoretical frameworks; framework development; framework validation	initiated (the evolutionary prototype)	review...Prototype 3: Field testing which led to prototype 4 of the checklist that is ready to be used in the English Foundation Programmes and which can be revised regularly.
Q 1. How can an appropriate method be established to design and develop an evaluation checklist for teaching materials in English language programmes?		
Q 2. What are the possible sources and basis for designing and developing teaching materials evaluation checklists?	Q 3. What are the design guidelines for the development of teaching materials evaluation checklists?	Q 4. How can teaching materials evaluation checklists be validated?

Table (9) General Processes of the Study

As is clear from the table, the main research question of this study is answered throughout the three phases of design-based research methodology and within the sub-questions 2, 3 and 4. Similarly, question two was answered within phase 1 where the main sources of the teaching materials evaluation checklist were identified through a literature review, setting needs and the developed conceptual framework. To answer question three, the researcher needed to reach phase two, before the design guidelines of the teaching materials evaluation checklist could be clearly identified. Finally, the answer for question four was achieved through the iterative cycles of formative reviews in phase three. All the procedures of the above phases were documented theoretically, procedurally and practically through the three phases. Van den Akker et. al. (2006: 38) advise that “all phases of the analysis process have to be documented, including the refining and refuting of conjectures” which will help “to ascertain the credibility of the analysis”. As a result, all the mentioned stages and processes documentation is presented within the thesis

chapters or in the appendices through participants' observational logs and written comments and feedback. The data collection processes and analysis are summarized next.

6.3 Understanding the collected data and their analysis

Each cycle of the formative review needed specific types of participants as their role may differ according to the goals of the assessment and the experience or the knowledge the participants have. Weston et. al. (1995: 35) suggested that there are “three different roles” of the participants within the formative review and these are the “evaluator, critic, or reviser”. In this study, and in each cycle of the formative review, the participants were encouraged to validate the developed checklist in any of the three roles. They could evaluate, criticize and revise its content, format and layout according to their own points of view or knowledge, as the overall purpose was to incorporate all the possible recommendations which will help in designing the teaching materials evaluation checklist as illustrated in chapter 5. Nieveen & Folmer (2013: 161) proposed that “design researchers need to select” the formative review “methods that fit the research questions”. In this study, the main research question purpose was to find the appropriate method to develop a teaching materials evaluation checklist. The answer for this investigation was provided through the three sub-questions that explored the through literature reviews to specify the possible sources and basis for designing and developing teaching materials evaluation checklists and the empirical phases (design and evaluation) that helped to yield the checklist prototype, the design guidelines and the method for teaching materials evaluation checklists validation. The researcher had to ensure that the instruments used, which were prepared in advance, were enough to collect all the needed data, and the number of participants in each cycle was manageable. Consequently, the participants and the methods were selected in accordance with the aims of the formative review and the types of users who are expected to use the teaching materials evaluation checklist. The expected users of the checklist are experts in teaching materials development and evaluation as well as teachers, coordinators and the educational institutions involved with English language programmes. Other users may include programme directors or other policy makers who might be also involved in teaching materials evaluation and selection. Almost all the expected users will need some sort of guidelines or support in using the checklist that can be in the form of an accompanying manual or heuristics for the teaching materials evaluation checklist design and use.

Each type of data was analyzed according to the purpose of their collection. For example, the data collected through brainwriting sessions were analyzed using a thematic analysis technique because that technique enabled the clustering of similar themes and categories together in order to use them in the teaching materials evaluation checklist design and construction. The data collected through formative review cycles were written in the form of a report after each cycle. Each report included detailed feedback from the participants, which was used directly after each cycle to make the recommended modifications to improve the checklist content and structure. This type of analysis is called 'sequential analysis', where the data "reanalysis" can be used "to ensure a more robust set of findings" (Miles and Huberman 2013: 175). During the sessions, the designer is usually present as an "observer" to gather review data and "(if necessary) to intervene if serious problems arise" (Tessmer 1993: 101). All the data collected were qualitative in nature and each cycle of formative review had its different preparation and methods for data collection. For the analysis of qualitative data, "there is no prescribed way to address the process." Researchers choose to analyse data using ways that "stem from a combination of factors, which include the research questions being asked, the theoretical foundation of the study, and the appropriateness of the technique for making sense of the data" (Kawulich 2004: 96). Consequently, data analysis after each cycle depended on the data sheets and tables that grouped the different comments of the participants in an easy and practical way in order to use them in the checklist improvement. Also, the research purposes and review purposes were explained to the participants and the review processes were managed well throughout all the cycles.

6.4 Teaching Materials Evaluation Checklist Rigour

To check the quality of the designed product or the checklist, the validity has to be established from the beginning and its robustness and rigour are clearly identified through its actual use by the potential users and usability testing. As Russell (2002: 2) specifies: "rigour, in terms of knowledge, arises out of our capacity and need to determine the actuality of actions as performances. That is, it is the "form" of "perform" that allows us to determine whether or not the state of can-do has been achieved". The assessment of the teaching materials evaluation checklist changed according to the aspects which needed to be tested. For example, "earlier alpha-style" reviews "tend to center on the internal structure of interventions (validity); during beta testing, use in context (practicality) receives more attention; and once interventions stabilize and are used under

representative circumstances, more robust gamma testing can take place (effectiveness)” (McKenney & Reeves 2013: 536). It is concluded from teaching materials evaluation checklist implementation, reviews and revisions that it has fulfilled the “three aspects of viability... practicality, relevance and sustainability” (Van den Akker et. al. 2006: 79). Reliability and validity are also dealt with by DBR researchers. The Design-Based Research Collective (2003: 7) stated that “objectivity, reliability, and validity are all necessary to make design-based research a scientifically sound enterprise, but these qualities are managed in noticeably different ways than in controlled experimentation”. According to Design-Based Research Collective (2003), a DBR researcher “typically triangulates multiple sources and kinds of data to connect intended and unintended outcomes to processes of enactment” so the “methods that document processes of enactment provide critical evidence to establish warrants for claims about why outcomes occurred”. Thus, reliability can be achieved through all the various phases and cycles of design-based research which involve several methods of data collections and the repetition that occurs to better the investigated interventions. Likewise, “validity of findings is often addressed by the partnerships and iteration typical of design-based research, which result in increasing alignment of theory, design, practice, and measurement over time” (The Design-Based Research Collective 2003: 7). Visscher-Voerman et. al., (1999: 16) divides the design and developments patterns in educational research into four paradigms: instrumental like instructional design models that are based on planning through using prepared objectives, communicative where the product is developed through negotiation between “developer, clients and users”, pragmatic where developer create their product and then test it with the users through the processes of “building, testing and revising several prototypes” and finally the artistic or the connoisseurship approach where there is no specific method for the product development and design. From the above approaches design-based research correlated to the (pragmatic approach), where the quality aspects are successful when “the design and evaluation activities are intertwined” and the “prototypes are regularly tested with users for their usefulness and effectiveness” Visscher-Voerman et. al., (1999: 24). It is also noticed that the design and development approaches are overlapped and difficult to separate. For example, the pragmatic approach that is used in design-based research can be instrumental, communicative and artistic as it may involve data collection activities using different techniques from these methods.

To fulfill the above quality conditions, great care was also taken when contacting and communicating with the participants. They were informed in advance of their roles and told that their participation was for the sake of validating and reviewing the checklist. Every instrument of data collection emphasized this at the beginning of all the sessions (one-to-one protocol as an example in (Appendix H1), and small group presentation in (Appendix J1). Participants were also asked to give their truthful comments, as this would help the researcher in improving the teaching materials evaluation checklist. This helped to avoid “the Hawthorne effect” which happens when the “people selected to try out a new product are flattered by the attention given them and reciprocate with less than honest or candid feedback so as not to offend the individuals who chose them to participate in the first place” (Beyer 1995: 57). The main criteria for assuring rigour in the checklist development, which were carefully considered are illustrated in Table (10). As confirmed by Nieveen & Folmer (2013: 151): “at the end of a design research project, the intervention should suffice all of these criteria. However, usually each iteration concentrates on one or two of these criteria.” Each of these criteria is discussed in the next sections.

Criterion	Requirements
Relevance (also referred to as content validity)	There is a need for the intervention and its design is based on state-of-the-art (scientific) knowledge.
Consistency (also referred to as construct validity)	The intervention is ‘logically’ designed.
Practicality	Expected The intervention is expected to be usable in the settings for which it has been designed and developed Actual The intervention is usable in the settings for which it has been designed and developed
Effectiveness	Expected Using the intervention is expected to result in desired outcomes. Actual Using the intervention results in desired outcomes.

Table (10) Criteria for high quality interventions (Nieveen 1999, in Plomp 2013: 29)

As statistical studies can never be used all the time as indicators of universal truths especially in complicated environments as the educational settings, practical alternatives are required to compensate for such methods. Yin (2003) differentiates between statistical generalization of survey studies and analytical generalizations of case studies. Unlike survey research that depends on statistical generalities for its external validity the investigator in case studies and design-based research “is striving to generalize a particular set of results to some broader theory (Yin 2003: 37). But in spite of that, “each context has unique characteristics that justifies that the design principles should be used as heuristic statements” and they “provide guidance and direction, but do not give ‘certainties’ (Plomp 2007: 24). Janet Schofield (2002, in Huberman & Miles) states that “most researchers writing on generalizability in the qualitative tradition agree that their rejection of generalizability as a search for broadly applicable laws is not a rejection of the idea that studies in one situation can be used to speak to or to help from a judgement about other situations” (Schofield 2002: 171). Moreover, Schofield (2007) indicates that “there is broad agreement that generalizability in the sense of producing laws that apply universally is not useful or obtainable goal for qualitative research”. Additionally, “current thinking on generalizability argues that thick descriptions...are vital” (Schofield 2007: 185). Thus, DBR can rely “on techniques used in other research paradigms, such as thick descriptive datasets, systematic analysis of data with carefully defined measures, and consensus building within the field around interpretations of data” (Design-Based Research Collective 2003: 7). Van Den Akker (1999: 12) suggests to “invest in ‘analytical’ forms of generalization” as readers can be supported to “make their own attempts to explore the potential transfer of the research findings to theoretical propositions in relation to their own context”. Schofield (2007) takes the argument further to state that “paying attention to where a phenomenon is in its life cycle does not guarantee that one can confidently predict how it will evolve” and that the “conclusions formed on the basis of a study conducted at one point in time will be unthinkingly and perhaps mistakenly generalized to other later points in time to which they may not apply” (Schofield 2007: 195). In that sense, generalizability in educational research becomes almost an illusion that can never been achieved unless it is conducted iteratively and over longer periods of time and with as many populations and settings as possible. It is concluded that this problem in educational research is better approached through DBR methodology, its various conceptions and iterations of the same problem. For example, generalizability aspect of this study lies in the developed conceptual framework in the

analysis and exploration phase. Actually, basing DBR studies on a model or a framework from the beginning does not only guide the researchers within the research phases, but it also reinforces the robustness of such studies through facilitating generalizability, future replications and linking educational studies where such frameworks can be used by other researchers instead of starting from scratch. Through this study it was assured that these quality conditions as validity, practicality and effectiveness were achieved in the developed teaching materials evaluation checklist as elucidated next.

6.4.1 The Checklist Validity

Validity of the checklist was checked several times throughout the formative review processes. It was the main focus of the first cycle through the two techniques (one to one and expert review) of data collection with their several instruments which included open-ended survey questions, observational logs and debriefing questions at the end of these sessions. The expert reviews were all directed towards checking both the content validity “state-of-the-art knowledge” and the internal consistency validity “coherence throughout the various...components” (McKenney & Reeves 2013: 539) of the checklist. In general, most opinions of the participants were positive in terms of the checklist validity as the early recommendations from the expert appraisal of the first prototype helped in shaping the content and format of the teaching materials evaluation checklist (Appendix G2). Throughout the different review cycles, there were only two instances where participants expressed their concerns regarding the checklist validity. The first was during the small group review, where one of the experts expressed his concerns about issues like the “face validity” of the checklist, especially that the checklist developer is an ordinary teacher who might not have enough experience to conduct such complicated task, where both academic and practical experience is required. Despite that, it is recognized that sometimes ordinary teachers are the appropriate investigators to do educational research as some researchers advocate. For example, Kincheloe (2012: 18) propose that “teachers must join the culture researchers if a new level or educational rigour and quality is ever to be achieved”. In that sense, the researcher had the opportunity to search for answers for many ‘hows’ and ‘whys’ without prior assumptions. So the researcher had to find the explicit and the implicit relations and connections between various disciplines through this study, a status that can be considered the key to theoretical and practical understanding of teaching materials evaluation. In other words, the researcher has a practitioner's background, besides approaching the topic through DBR methodology

which enabled incorporating both practical and theoretical knowledge in developing a comprehensive and clear checklist for potential users.

6.4.2 The Checklist Practicality

The assessment and review of the checklist practicality was accomplished through two main cycles: small group reviews and field testing. Practicality here is related to the usability of the teaching materials evaluation checklist in the real settings where it is expected to be used. The two aspects that received the most dissatisfaction were the length of the checklist and the rating scale. As one of the experts stated “the overall length of the checklist, combined with the somewhat opaque rating scale, discourages or hinders the probability of a completed checklist” (Expert 1, small group review cycle as explained in chapter 5). As a result, most of the adjustments were related to those two parts of the checklist. The rating scale was changed three times. Also the items that were not essential in the process of teaching materials evaluation or repeated in other sections were removed. In the field testing sessions, the usability problems were fewer compared to the other cycles. This is due to the iterative assessment of the checklist through the formative review data collection instruments and the useful feedback received from the experts, the teachers and the coordinators who participated in these sessions.

6.4.3 The Checklist Clarity and Appeal

Most of the issues relating to the clarity and appeal of the checklist were noticed through the researcher’s observations. These observations enabled the researcher to discover the interesting parts of the checklist as well as the difficult and confusing sections through the different sessions. The users appreciated the checklist comprehensiveness and liked the section about “what to look for” as this section provides clear items and details on what to search for in the selected or evaluated teaching materials. There were some suggestions to make the checklist appealing, such as providing background information in order to make sure that the checklist will be used appropriately, and some kind of training or payment to encourage the use of the checklist among teachers (one-to-one reviews, Teacher 1). A few comments were also made concerning the length of the checklist and the sub-headings, and the recommendation was “to use as less items as possible” (Teacher 1, small group review cycle). There were also suggestions to increase the appeal of the checklist by changing the general layout to a more familiar one, that the teachers are used to when dealing with teaching materials evaluation: “the checklist needs to organize items according to their categories” (Teacher 4, field testing) which

was explained in details in chapter five. All the formative review cycles and sessions helped to discover the unclear items and the less-friendly parts of the checklist and in the subsequent versions of the checklist they were revised and corrected to make the checklist more interesting and attractive for its potential users.

6.4.4 The Checklist Effectiveness

The checklist effectiveness here was achieved in the last cycle of the formative review when the use of the checklist in the six Colleges of Applied Sciences proved to be useful. It achieved the expected results for evaluating the teaching materials in the English Foundation Programmes when used by ten different participants. Beyer (1995) specified that effectiveness is also investigated in almost all of the formative review sessions. In answering the question “what is effective?” he proposed that it is “what users find particularly clear, helpful, easy to use, efficient or interesting” (Beyer 1995: 26). Plomp (2013: 28) distinguishes between “expected and actual practicality and effectiveness” suggesting that “only when target users have had the opportunity to use the intervention in the target setting, the evaluator will get data on the actual effectiveness.” Therefore, effectiveness is mostly noticed through the field testing cycle where the participants were able to use the checklist in evaluating the teaching materials that they had taught or they were teaching during the different sessions of that formative review cycle. During these sessions, the researcher aims were to check the time spent using the checklist, the behavior of the users and the problems they encountered while using the checklist. In general, the ten users of the checklist were able to complete the evaluation of their textbooks within a reasonable amount of time and without any huge usability problems. After these sessions, the teaching materials checklist went through its final version review that is intended to be diffused later in the Colleges of Applied Sciences as a validated instrument that can be used in teaching materials selection and evaluation in the English Foundation Programmes. However, even after its distribution and use, the checklist is still liable to more improvement whenever this is needed. This is one of the advantages of using DBR in designing educational products where the opportunity for continuous improvement is encouraged and supported.

6.5 Triangulation within the Study

As is clear from the above sections, the formative review cycles enabled the triangulation of data collection instruments, participants and settings. The assessment process went through four major iterations which basically represented the formative review cycles.

Miles & Huberman (1994: 438) referred to triangulation as “the term most often used in connection with analysis and confirmation issues...a term with multiple meanings”. So the meaning of triangulation that will be used here is the one that means “the study of the same phenomenon through applying and combining several data sources, research methods, investigators, and theoretical scheme” (Wang & Duffy 2009: 275). Design-based research, combined with the use of several methods such as formative review, allows for automatic triangulation on several levels: triangulation of information through consulting various sources and literature reviews, triangulation of methods of data collection, the participants and different settings. All of these multiplicities in conducting research strengthen the results of study and as Nieveen & Folmer (2007: 163) suggest “the effectiveness of triangulation rests on the premise that the weaknesses in each single data source will be compensated by the counterbalancing strength of another.” Even with a single cycle of formative review, triangulation occurs through using different instruments of data collections and from involving different participants and settings.

6.6 Developing Teaching Materials Evaluation Checklists: the Results

Richey & Klein (2007: 129) specified which results are considered valid and reliable of any DBR project. These include expanding “the knowledge base” leading “to new research” and establishing “the foundations of new theory”. Similarly, Kelly, Lesh & Baek (2008: 322) emphasized that any DBR study must include and specify the “designed product, the context within which the design was implemented, warrants for evidence of the local impact, the theoretical assertions and their relations to the design work, the conditions through which the theoretical assertions are generated and warrants for evidence of the theoretical assertions”. It becomes clear that in DBR, knowledge is not confined to theoretical aspects, but also includes the practical instruments and products that are created through its phases and which can be achieved within one study. Wilson (2014: 2-3) states that “knowledge *is distributed* [italic in the source] in different ways and combinations, and knowledge is always ‘stored’, used, and transmitted within a complex web of interrelated resources and people” so, knowledge is not simply in the people’s minds, but is also exhibited in “the artifacts and practices built up over time.”

To reach the above results, the teaching materials evaluation checklist development is based on theoretical and practical phases of design-based research. The analysis and the exploration phase resulted in the development of the underpinning tenets of the checklist sources and basis that were converted into the conceptual framework. Through the design

and construction phase, the checklist prototype was developed whereas the evaluation and reflection phase facilitated the review and assessment of the teaching materials evaluation checklist prototype with actual users as well as the researcher's reflections. McKenney and Reeves indicate that reflection in design-based research "involves active and thoughtful consideration of what has come together in both research and development (including theoretical inputs, empirical findings and subjective reactions) with the aim of producing new (theoretical) understanding" (McKenney & Reeves 2014: 150). All of these processes contributed to the development of different insights about the checklist design, its review and use in the English Foundation Programmes. One of the main results of this study is the constructing of design guidelines on to how to develop a teaching materials evaluation checklist. Throughout the formative review cycles not only the problems and strengths of the checklist were discovered, but also the appropriate procedures for designing materials evaluation checklists were identified. These guidelines are very important for several stakeholders and different developers. They are important for research because "these principles show the contribution of design research to the existing knowledge" and "for educational designers, these principles carry rich information on how to design similar interventions for similar settings". For the "future users, the principles provide information needed for selecting and applying interventions in the specific target situation and provide insights in the required implementation conditions" and finally "for policy makers, these principles assist in making research-based decisions for solving complex educational problems" (Nieveen and Folmer 2013: 154). Below are the eleven general guidelines on how to develop teaching materials evaluation checklists for the English Language Foundation Programmes. These guidelines are clear and straightforward and they can guide the checklist developers through simple and easy instructions.

- 1) Construct an outline for the quick teaching materials evaluation checklist; it can be based on any impressionistic or first glance checklist recommended by researchers such as McGrath (2002).
- 2) Think about the two main sources for the detailed teaching materials evaluation checklist design: research in the field and local context and needs.
- 3) Select the basis or the sources for the checklist development (research, setting needs or both).
- 4) Construct a conceptual framework for the checklist design and test it with subject matter experts

- 5) Select the main categories and items of the evaluation checklist based on the main categories of the conceptual framework
- 6) Create an initial prototype (s).
- 7) Integrate the developed prototypes (based on the conceptual framework) into one comprehensive prototype that incorporates significant categories and related items.
- 8) Evaluate the developed prototype to check validity, clarity, practicality and effectiveness using formative review methods (expert review, one-to-one review, small group review and field testing)
- 9) Redesign the checklist based on the feedback after each formative review cycle.
- 10) Make the final appropriate revisions on the developed checklist.
- 11) Update the developed checklist whenever needed.

These general guidelines are the fourth result of this study. The first result was specifying the teaching materials evaluation checklists sources through a validated framework; the second was the evaluation checklist and the third was identifying the validation method for such instruments which is formative review technique. As any developer is supposed to “recommend the way in which products and tools should be used” (Richey & Klein 2007: 134), guidelines on how to use the checklist were also formed. So the following are the instructions of use for the developed teaching materials evaluation checklist as explained to the users in the field testing cycle:

The checklist is divided into two main parts. Part A: quick evaluation and part B: close or detailed evaluation. Each part is on a separate sheet for ease of use.

- 1) Go through the checklist to become familiar with its format and content.
- 2) Bring the materials to be selected or evaluated.
- 3) Evaluate the materials against the items in the quick evaluation checklist first.
- 4) If the evaluation score is above 80% go to the detailed evaluation sheet.
- 5) Repeat the evaluation for all the components in the detailed evaluation sheet by selecting the appropriate answer from the drop-down list.
- 6) If the total score is 60% or above the materials can be selected or reused for students and if it is less, discard and look for other appropriate materials.

The validation of the teaching materials evaluation checklist is also recommended within the results of this study. It is thought to be an easy, clear, practical and useful review tool

that enables the developer to interact with the actual users and stakeholders, as well as providing a systematic assessment tool for regular improvements. These steps are as follows:

- 1) After the construction of the checklist first prototype, screen it against the available well-known practical checklists [researcher suggestions: Tomlinson (2013), Wilson (2013), Bichelmeyer (2003), and Stufflebeam (2000)].
- 2) Discuss/ review it with actual users in any appropriate form (e.g. one-to-one/ focus groups/ small group reviews).
- 3) Discuss/ review it with a group of experts and then a group of teachers or potential users
- 4) Ask users (experts and teachers) to use the evaluation checklist in evaluating their teaching materials
- 5) Adjust and amend the checklist according to the feedback from the users (experts/ teachers)

The evaluation checklist is a multiple-roles-evaluation method ensuring quick and reliable evaluations, and it empowers the professional development of all the involved stakeholders. Harwood (2010:8) indicated that materials evaluation results can yield “useful messages” even for the teachers “who are producing or adapting in-house materials.” The professional development element of materials evaluation has arisen in almost all the cycles of formative evaluation, as was noticed in cycle 1, expert review and small group review cycle 3, in experts’ sessions, and through the impressions of different participants who took part in this study. So teachers’ involvement in constant evaluation can be the solution for teachers’ professional development, the understanding of their learners’ need and the improvement of the whole programme. One of the researcher’s inferences from observations of the teachers’ use of the checklist in evaluating teaching materials is that teachers who were involved in PhD courses or who have conducted some sort of materials evaluation during their training courses were more thoughtful, critical and aware of the courses that they are teaching. As the first option of pursuing higher education is not available for most teachers and practitioners in these English Foundation Programmes, providing courses and workshops on materials evaluation can be the appropriate alternative. Moreover, in ELT settings, teacher training is not usually provided and though the “alignment” of curriculum “with pre-and in-service teacher education is critical to successful implementation” (Van den Akker et al.2006: 70), teachers in English Foundation Programmes are excluded from the engagement in materials development, selection and evaluation.

Through this study, several interesting results were achieved within its different phases besides the above mentioned practical results. The teaching materials evaluation checklist turned out to be a multi-purpose instrument not only an evaluation tool. In fact, Richards (1993: 10-12) uses “designing criteria for evaluating textbooks” along with “examining the content of textbooks” and “trying out materials’ design” as instruments for his teachers’ training workshops. Also Henrichsen (1983: 23), in his international survey results with 153 teachers’ participants, specified the top ten teachers’ needs and the “training in TESL materials selection and evaluation” was ranked number 2 among other needs as “specific training in how to teach” and the language skills (listening, reading, writing and speaking). As the checklist is considered here a professional development tool, so the headings and sub-headings and the items of the content of the checklist become very important. It was thus concluded that when teachers or evaluators encounter headings such as ‘SLA’, ‘ELT curriculum’, ‘teaching principles’ and ‘setting needs’, they will acquire new knowledge on all the mentioned fields and areas instead of using the previous traditional categories and organizations of checklists development such as organization, content, skills, tasks, which may aggravate practitioners' sedentary routine rather than provoking criticality and professional development. If a teaching materials evaluation checklist is considered as a comprehensive professional development instrument, or even a training course for teachers in English language programmes, then materials evaluation is clearly one of the top priorities for authorities in English Foundation Programmes in Colleges of Applied Sciences. As mentioned in previous sections, the uses of the teaching materials evaluation checklist are numerous. Also, it was recognized that providing a short summary about the design and the use of the checklist can facilitate the evaluation processes and encourage teachers and other stakeholders to use it. Similarly, it was discovered that evaluation in itself can be used as a training tool in English language programmes for all stakeholders, especially teachers, which offers them the needed background and experience to improve their courses. Within this study, ten uses from the theoretical and practical processes and phases are identified. The uses revealed through research showed that teaching materials evaluation checklists can be used to:

- 1) Evaluate textbooks while-use.
- 2) Evaluate new textbooks.
- 3) Select a textbook series from different packages or copies

- 4) Improve the materials as a post-use evaluation in order to provide suitable add ones or ancillaries.

Whereas the uses discovered through the practical aspect of this study through the review and use of teaching materials evaluation checklist include its use as:

- 5) A tool to help teachers and writers to develop their own teaching materials.
- 6) A professional development tool (on personal and institutional levels)
- 7) An educational training tool (on institutional level)
- 8) A knowledge assessment tool (to assess teachers' experience)
- 9) A Link that connect research findings in materials development and evaluation to practice through its content and regular use in English Foundation Programmes.
- 10) A Quality management tool in English language programmes

In addition to the above theoretical and practical results, findings based on comparing literature reviews and data collection revealed certain theoretical and practical characteristics about the teaching materials and their evaluation in the English Language Foundation Programmes. The main conclusions are summarized in the table (11) below.

Study Findings based on the literature and data collected for the teaching materials evaluation checklist main categories

The selected areas from literature	
SLA Principles	Teaching Principles
ELT Curriculum Design Principles	
<p>The three areas (SLA principles, Teaching principles and ELT curriculum design principles) are connected in theory and practice as it can be noticed in the items of the teaching materials evaluation checklist. This could be attributed to the influence of SLA studies and the curriculum studies and the “combination of three paradigms cognitivism, constructionism and humanism through learner-centered education” Reigeluth, Beatty, & Myers (2017). The integration of paradigms also led to approaches that focus on common themes from various fields such as eclecticism which involves “picking and choosing some procedures from one methodology, some techniques from another, and some exercise formats from yet another.” Tarone and Yule 1989 cited in Griffiths, 2008).</p>	
<p><i>Literature (Theory, empirical, practical)</i></p>	<p>In the ELT curriculum design, the changes in pedagogy “had led to a shift of interest from research on teaching to research on learning” Stern (1989: 207). As a result the educational research began to focus on classroom settings from two aspects the pedagogic and the linguistic</p>

points of view. As the two conceptions “have not been sufficiently distinguished” many issues are found in the three areas which lead to confusion among researchers and practitioners because of the “constant shift from talking in linguistic terms to talking about pedagogy” in the literature as Johnson (:1989: 209) explains. Reigeluth, Beatty, & Myers (2017) propose five “educational principles or guidelines for learner-centered education” where achievement is based on learning rather than time, instruction based on authentic tasks that are personalized to different students, which require the transformation of the roles of learners, teachers, curriculum and technology in the educational processes. Most of these principles are reflected in the literature and practical needs analysis in this study using various terms and conceptions.

The selected areas for setting needs

Learners needs	Teachers needs	Institutional needs
-----------------------	-----------------------	----------------------------

Data collected

The learners’ needs revolved around almost the same areas in this study and many other similar studies. These are the incorporation of all the skills, exploiting the available sources like internet, mobile phones, newspapers and magazines sufficient activities of dictionary use, dictation, and spelling games, illustrative drawings, pictures, maps and infographics, entertaining activities like songs, short films and documentaries and allocating enough time for students to speak in class and be involved in teaching and presenting parts of the lessons to the class.

Teachers’ criteria for textbooks quality and selection are mostly related to their students’ needs and their own daily concerns and problems when teaching general English textbooks as well as their general knowledge about the ELT research. Teachers’ needs have changed and instead of focusing on teachability aspects in teaching materials, students’ needs and teachers support and professional development are the issues that need more emphasis in the English language programmes.

The institutional needs in Oman -like most developing counties needs-depend on “the role of English in the country” the “sociocultural environment” and “the type of tested used” McDonough, Shaw and Masuhara (2013: 8-9). Thus, these issues are important in materials development and evaluation.

Summary of Findings

The practical results of the study correspond to the theoretical principles of leaner-centered education.	Teaching materials main role is to facilitate the teachers’ role to help their learners acquire the language. Teachers supports in form of professional	Institutional needs and policies in Oman are mostly connected to the international influences and standards as well as local restrictions when selecting and evaluating
---	---	---

development is a must (Roblin, Schunn & McKenney, 2018) ELT curriculum and materials.

Table (11) Summary of Theoretical, Empirical and Practical Findings

Through the various processes and phases, it was concluded that the checklist can be perfectly used in groups with a short presentation or an accompanying manual or website. It can be also used on an individual basis when the evaluators have enough experience and background about teaching materials development and evaluation. When used in the English language programmes, the teaching materials evaluation checklist can fulfill the above mentioned uses and it can help in selecting, improving and comparing the different English language teaching materials. Thus, the authorities should emphasize and focus on the evaluation of the teaching materials to assure the success of their English Foundation Programmes. Colleges and universities are also advised to incorporate the analysis and the evaluation of teaching materials in the teachers' training programmes through several semesters, not just one introductory module or an elective.

Spreading the results of this study will have short-term goals and long-term ones. The initial aim is to provide two forms for the teaching materials evaluation checklist: an electronic version and hard copy, with specific guidelines on how to use them in evaluating teaching materials in the English Foundation programmes. In the electronic version, each "evaluation question" will "be given a numerical value" and "total scores can be calculated and indications can be derived of the potential value of the materials" (Tomlinson 2003: 16). Each part of the checklist will have a cutoff point score where the results of the evaluation should not fall behind 80% for the quick checklist and 60% for the detailed checklist. The printed version will be the same, but the calculations will be done manually by the evaluator. The initial dissemination will include (1) the checklist for materials evaluation and (2) a manual of the evaluation checklist development and design and how it can be used so that other developers and evaluators can benefit from it. The checklist diffusion will hopefully be through a designed website which will include what teaching materials evaluators need to know in order to develop their own checklists and conduct their materials evaluations. Through this website, the checklist will be available for all users at any time along with the availability of all the information they need in order to develop, use and validate materials evaluation checklists.

6.7 Developing Teaching Materials Evaluation Checklists: Reflections

The overall aim of this study was to discover an appropriate method to produce a viable checklist to help developers and users, especially teachers, coordinators and institutions, to find out the basis or sources for teaching materials evaluation checklists. In other words, it aimed to discover how teaching materials evaluation checklists are developed used and validated. From the research point of view, it is hoped that the study contributes to both the field of materials evaluation and of design-based research methodology. The designed checklist has helped teachers and potential users to think about what they are teaching and how they can approach the teaching materials they are using every day. This means that regular materials evaluation and the use of the evaluation checklist in particular can guide teachers on how they select, use and evaluate their teaching materials. The evaluation process can also improve the English Foundation Programmes and can help educational institutions to achieve their stated goals and missions.

The implications and results of this study go beyond the refinement of the teaching materials evaluation checklist. As Richey & Klein (2007: 125) suggest: “perhaps one of the hardest parts of the researchers' job is to make sense of their findings in such a way that enables others to learn from their experiences”. What also makes it more complicated in design-based research methodology, is the researcher’s attempt to keep a balance between the participants’ assessment of the instrument and the conclusions the researcher can deduce which accompany such reviews. This problem is solved through separating the two aspects in the third phase (evaluation and reflection). By doing that, the evaluation process is used to test the developed instrument by the actual users, and the reflection process is used to enable the researcher to express and explain the general conclusions about the study as a whole.

Besides the main upshots in this study that are demonstrated in section 6.6 (based on phases one, two and the first part of phase three), five other themes or dimensions can be identified (from the literature and the study practical processes), related to designing and developing teaching materials evaluation checklists. These themes were based on the reflections that occurred towards the end of this thesis which McKenney & Reeves (2014) call “organic reflection” or the “intended contemplation”, where techniques like “well-timed breaks” were used when thinking of the teaching materials evaluation instruments as a whole from different aspects and after answering all the research questions. According to McKenney & Reeves (2014: 148) “reflection involves active and thoughtful

consideration of what has come together in both research and development (including theoretical inputs, empirical findings and subjective reactions) with the aim of producing new (theoretical) understanding”. These themes should be of help to other developers and users, who are interested in teaching materials evaluation instruments, which will enable them to understand their development and use. It was noticed that these issues have been dealt with by several researchers separately, which has resulted in isolated results and interpretations and which impedes establishing a complete representation and view about the teaching materials evaluation checklists as a unified whole rather than scattered pieces. These intersected issues that are dealt with separately make the efforts of different researchers unclear and incomplete, which leads to a neglect of the role and importance of evaluation in education as well as its different instruments design. Consequently, this has weakened the focus on an important and vital aspect of English Foundation Programmes (materials evaluation) which in turn has resulted in lack of awareness on the development of evaluation instruments and their different uses.

Thus before designing or developing any teaching materials evaluation checklist, the developer is advised to recognize and to work on the five dimensions: (i) reviewing previous checklists to know their problems, (ii) investigating their design frameworks and schemes, (iii) looking at how these checklists are used in evaluating teaching materials, (iv) considering how they are validated and finally (v) reviewing studies that examine their usefulness in English language programmes. In other words, the first theme that emerged by the end of this study is the investigation of previous evaluation checklists and the identification of their advantages and disadvantages. The aim should not be to use them as a base for the new developed checklists, but rather to see how these teaching materials evaluation checklists are developed. The second theme is the search for any frameworks used to design these checklists and decisions about the content and processes of the checklist development, including the need for design guidelines and rating scales. The third theme is a consideration of how these checklists are used in conducting teaching materials evaluations and how to interpret and use the results of these evaluations. The fourth theme is validation of the checklist with both experts and intended users and decisions on the method that can be used to validate and review these instruments. The last theme is about its usefulness when used by different users and an acknowledgment of the evaluation checklists as multi-uses instruments in the English language

programmes. Each one of these issues or dimensions are important for any evaluation checklist developer. These dimensions will be explored in detail next.

In chapters one and two the need for the development of an evaluation instrument for teaching materials in the English Foundation Programmes was discussed, as well as the previous evaluation tools designed by other researchers and developers. The re-discussion of this matter here is for the sake of clarifying the above themes about teaching materials evaluation checklists. As there are many evaluation checklists and schemes for teaching materials evaluation, any new developer or user can be lost among such an overabundance of instruments. So considering such themes from the beginning will help designers to save time, to organize their ideas and narrow their focus. In fact, the first and second themes are achieved through the review of the existing teaching materials evaluation checklists and the available frameworks. Some models and criteria are represented in chapter one in section 2.8 and some will be also discussed in this section as samples for themes (i) & (ii) of the reflections results. For example, Daoud & Celce-Murcia's (1979) checklist has two sections, one for the textbooks and the second for the teachers' manuals with yes/ no questions and a five point rating scale. In this checklist, there are no explanations of the items nor the use of the rating scale. Also, there are more than one item within one question. Byrd & Celce-Murcia (2001) has three sections: the fit between textbook and curriculum, between textbook & students and textbook & teachers, but it is too general and also with no clear explanations within the checklist. Littlejohn's (2011) three levels framework - which explores (i) what is there: publication, access design, (ii) what is required of users: subject matter, activities, participation, and (iii) what is implied as: aims, selection, sequencing roles of learners and teachers. His framework is also too complicated and ambiguous for users, especially practitioners. His framework is also too general, though it is based on matching the two aspects (situation analysis and materials analysis) to the "target situation of use" (Littlejohn 2011: 203-204). Griffiths' (1995) checklist has twelve criteria and is also too general and without a rating scale, and the explanations for each criterion are too long and general. Tucker (1975) has a two-section checklist: internal (pronunciation grammar and content) and external. The external part includes some vague and unclear items that cannot be evaluated as "competence of the author", as this criterion cannot guarantee the quality of the textbooks. Cunningsworth's (1975) checklist is inclusive but very long and difficult to use by practitioners, as it includes many terms and obsolete items such as methodology.

It also mixes analysis and evaluation questions in different categories and headings. The ELT Document published by the British Council in 1987 includes several evaluation criteria such as those produced by Cunningsworth (1987), Breen and Candlin (1987), Dougill (1987) and Hutchinson (1987). These attempts have also their drawbacks despite their pioneering roles in developing frameworks and checklists for materials' evaluation. For example, Breen and Candlin's (1987) designers' guide in two phases, initial questions about aims, content and users and phase two about learners' needs and classroom situation, is very thought-provoking, but lacks instructions for practical application. Sheldon's (1988) fourteen criteria are clearer than the others mentioned above as he provides an explanation for each criterion on the relevant subsequent page; but the rating scale is not practical as it combines both score system and comments in the same column for evaluating different items. Roberts' (1996) scheme to clarify materials evaluation tried to cover the evaluation process from pre-publication (piloting, decision stage 1) to post-publication (decision stage 2, classroom trial, summative evaluation). But his checklist is not really "demystifying" materials evaluation as it makes evaluation more difficult and challenging. Roberts Arthur (1980) checklist for selecting and evaluating social studies materials is also too general because it is meant to be used with all social studies subjects. It includes also a short paragraph at the beginning on its use and rating scale but with different values for different sections, which complicates the evaluation processes. It also mixes quantitative and qualitative questions. Williams' (1983: 251) checklist is based on four assumptions - up-to-date methodology, guidance for non-native teachers, needs of second language learners and relevance to socio-cultural environment - evaluated against a "set of linguistic, pedagogical, general and technical criteria" with no rating scale or clear guidelines. Ansary and Babaii (2002), who tried to locate the universal characteristic for EFL/ ESL textbooks evaluation, were unsuccessful in achieving that goal for several reasons. First, their work is based on limited sources and reviews (ten textbooks and ten previous checklists). Second, their terms and headings do not facilitate the use of the checklist by different stakeholders, as the criteria are very general and confusing for some evaluators (e.g. approach: the nature of language, the nature of learning and how the theory can be put to applied use) (Ansary and Babaii 2002: 6). Garinger (2002) seems to present rather an analysis checklist, which is not suitable for evaluation purposes and which also lacks clear instructions for use and a rating scale. Shaterly & Azargoon's (2012) checklist, with 16 yes/ no analysis questions, shares the same problem with Garinger 's checklist. It is a mix between a previous checklist by Mickley (2005) and new added

items suggested by general English professors. AbdelWahab's (2013) four-page teaching materials checklist is a comprehensive checklist but with an unclear rating scale, which makes its use imprecise. The most amended scheme has been developed by Mukundan. It started in 2009, when Mukundan presented a 'composite framework' where he tried to triangulate the development process through the use of more than one framework. While the use of multiple frameworks for the evaluation instrument may not be the solution, Mukundan's (2009) methodology of design and validation is one that will make huge differences to work in checklist practicality and effectiveness. But, his framework (based on Skierso 1991), with the supplementary tools or frameworks (concordance software and reflective journal), are complicating the process rather than simplifying it. Despite these efforts, what is lacking is the original basis for the checklist formation. The procedures include Mukundan et. al.'s (2011) tentative checklist with two sections - general attributes and learning-teaching content -, Mukundan et. al.'s (2011) modified version, where some items were inserted and others were deleted, Mukundan & Nimehchisalem's (2012) revised version, with two items removed, Mukundan & Kalajahi's (2013) use of the developed checklist to evaluate textbooks by teachers in different years for modification, and Nimehchisalem & Mukundan's (2015) last revision with comments of a panel of experts. Throughout the six-year process, "a qualitative method was used to collect and analyse the data. The checklist was refined, based on the experts' comments, problematic items were removed or revised and a scoring guide was added to it" (Nimehchisalem & Mukundan 2015: 789). Despite all of these processes, tricky items such as "the book has a nice feel" can still be found in the checklist. This criterion may affect the evaluation results because the nice colourful layout, for example, is not a reliable criterion for judging the teaching materials and can be problematic for different teachers.

The third theme here is to look at how some of these checklists are used in evaluating teaching materials in the English language programmes. The designer can create a folder that includes the checklists that were used in evaluating teaching materials in the English language programmes. This step will help to understand and realize the useful features that help in successful evaluations. For example Ranalli (2002) used "the four-guideline approach proposed by Cunningsworth (1995) to evaluate the New Headway Upper-Intermediate textbook, "one of the coursebooks used at the Foreign Language Institute of Yonsei University in Seoul, Korea" (Ranalli 2002: 2). This evaluation led to recognizing

some of the textbook's disadvantages such as “the methodology, which is somewhat restrictive and rests on some arguably shaky theoretical foundations” where the “approach to accuracy work is rule-based and behaviorist” and which “can be overcome through adaptation and supplementation” as the “book’s faults are outweighed by its strengths” (Ranalli 2002:17). Rahimpour’s (2013) evaluation of Top Notch 2 led to a suggestion of “the inclusion of more consciousness-raising activities, genuine negotiation of meaning tasks, and effective cooperative learning strategies would have improved this particular aspect of the book” (Rahimpour 2013: 771-772). White (2001) applied the framework designed by McDonough and Shaw to evaluate the High Impact series, which allowed him to gain a “more thorough understanding” of these series, which will help, along with his “knowledge of retrospective classroom implementation”, to meet the learning needs of his students (White 2001: 17). Nahrkhalaji (2012) stated that the use of a developed checklist “can help ELT teachers to make decisions about adaptation and adoption of the materials” (Nahrkhalaji 2012: 184). Also, Hamidi et.al (2015) and Hamidi et. al. (2016) used Daoud and Celce-Murcia’s (1979) checklist to compare two general English textbooks: Four Corners 1 and Top Notch Fundamentals in the first study and New Interchange 2 and Four Corners 3 in the second. They concluded that “curriculum developers, syllabus designers, and EFL teachers may find the findings useful in their language teaching practice” (Hamidi et.al. 2015: 1192) as well as selecting the appropriate textbooks. Hamidi and others also concluded in another study that, in some instances, “Four Corners 3 was found to be better than New Interchange 2” (Hamidi et. al. 2016: 2). Çakit (2006) evaluated New Bridge to Success 3 from the perspectives of students and teachers, which helped in engaging the stakeholders in teaching materials evaluation and in informing the evaluators about the strengths and weaknesses of teaching materials. Griffie & Gorsuch (2016: 6), through their investigation about what teachers know and do not know about their daily routine teaching, concluded that teachers may “know that some aspects” of the course “are working” and “know (or at least suspect) that some are not”. This conclusion can be generalized to almost all teachers especially if they are not involved in the course planning and development. In such a perplexing situation, involvement in teaching materials and courses evaluation can help teachers overcome such ambivalent thoughts and interpretations. Other studies in the literature that involve investigating teachers’ perspectives about teaching materials include Alemi & Sadehvandi (2012), Riasati & Zare (2010) and Ahmad & Shah (2014). Also, examples of studies that consider students' perspectives are Alavinia & Siyadat (2013) and

Hanafiyeh & Koosha (2014). Alavinia & Siyadat (2013: 154), for instance, used a “textbook Evaluation Tool (TET)” that was based on “Cunningsworth (1995) checklist”. This study investigated students’ “opinions on various aspects of four English textbooks, i.e. American English File 1, American Cutting Edge 1, American Headway 1, and New Interchange 1” concluding that “through using a convenient website related to a specific course book, the learners would be highly motivated and enjoy the English language learning process” (Alavinia & Siyadat 2013: 150). Other famous studies to compare and evaluate different textbooks include Masuhara et. al. (2008), Tomlinson et .al. (2001) and Tomlinson & Masuhara (2013). Their evaluations are based on their own developed criteria, which are based on research and SLA principles. Thus, by looking at how the checklists are used in evaluating the teaching materials and the results of these evaluations, the developers can understand the components that need to be included and the benefits expected from designing such evaluation instruments.

The fourth theme or dimension that a teaching materials evaluation checklist designer should consider from the beginning is the validation of such instruments. In fact, “the validation of a checklist plays an important role in establishing the credibility and utility of the checklist—particularly when the checklist is used for evaluative purposes” (Martz 2010: 222). Martz used mixed methods, “survey research and case study research”, to validate the developed instrument. This theme can help the developers to think of the appropriate method for testing and reviewing their developed instruments. In fact, developing an educational instrument or process without considering how it will be assessed and reviewed is like designing a product without providing the maintenance manual.

The fifth and final theme is a consideration of the studies that have investigated the usefulness of teaching materials evaluation checklists with potential users. A clear example of this theme is Nimehchisalem & Mukundan's (2013) method, who used a questionnaire based on a “modified version of an instrument developed to evaluate the usefulness of a writing scale” (Nimehchisalem & Mukundan 2013: 697) to test their checklist. As they mention: “one of the main limitations of this study is that it only considers the views and perceptions of a group of English language teachers” (Nimehchisalem & Mukundan 2013: 810), which means that teaching materials evaluation checklists developers should consider testing their instruments usefulness with all prospective users. As is clear, studies conducted in validating the checklists while

development and testing their usefulness after development are scarce. This could be attributed to two reasons. First, the absence of a methodology that enables the developers to consider all these dimensions at the same time in one study. Second, the lack of time and funds to conduct such studies. The suggested solutions through this study can be through using DBR methodology for designing teaching materials evaluation checklists and many other educational products and processes.

From the above discussion of the various schemes and criteria of teaching materials evaluation checklists, it is concluded that there is a huge production of evaluation instruments, but that their impact and use in educational contexts does not meet expectations. In fact, “the process of selecting an appropriate text has not become any easier for most teachers and administrators” as the process of evaluating textbooks is conducted inappropriately: some educators may “ask so many questions” and “others choose a reading textbook with little or no evaluation” despite the fact that a textbook will become “the centerpiece of the curriculum until another haphazardly chosen reader replaces it” (Wen-Cheng et. al. 2011: 91). In order to offer an evaluation instrument that can have more impact, the whole design process that identifies the checklist sources has to be clear, as well as the instructions for use, a method for regular review of the teaching materials evaluation checklist and a clarification of its vital role and uses in the English Foundation Programmes. The evaluation checklist developed through this study has tried to consider all of these issues through DBR which enabled the researcher to explore all the different aspects related to the checklist development.

6.8 Design-Based Research and Criteria of Success: Reflections

The results of the multiple phases and trajectories of DBR are multifarious as illustrated above. In fact, the purpose of design-based research exceeds artifact design and goes “beyond just creating designs that are effective and that can sometimes be affected by tinkering to perfection,” because “a design theory explains why designs work and suggests how they may be adapted to new circumstances.” As a result and “like other methodologies, design experiments are crucibles for the generation and testing of theory” Cobb et al. (2003: 9). McKenney and Reeves (2012: 31) define theories as “explanations of real world phenomenon substantiated by scientific evidence. They provide models or schemes for understanding the nature and causes of certain phenomena”. In design-based research the practical results in the form of the developed instruments and design guidelines are considered descriptive theories which “describe real world phenomena”

and that are “derived from empirical observation” whereas “theoretical understanding is developed through reflection...and especially through reasoning” which is “a rational thought process by which existing ideas give rise to new ones” (McKenney and Reeves 2012: 32). Besides descriptive theories, “DBR aspires to produce explanatory accounts that are not solely descriptive” so “theory in DBR is closely related to practice, and this link has its roots in the origins of the approach” Reimann (2011: 39). Thus as McKenney and Reeves (2012: 37) point out “the explanatory and predictive power of theory is especially needed to design interventions that solve real problems; and theories that serve normative prescriptive purposes are required to transplant and refine interventions”. As design-based research combines both research and design, the resulted theories or conjectures usually embrace the four elements of any theory mentioned by Whetten (1989 cited in Friedman, 2003: 516) which answer six questions: “(1)what, (2)how, (3)why and (4) who-where-then”. For example, the developed checklist in this study can be assessed with criteria of judging ‘what’ to ensure its “comprehensiveness and parsimony”, the ‘how’ criteria which refer to the relations between the factors and items of the checklist that were identified in answering ‘what’. The ‘why’ criteria is about the justification of selection of the specified parts, sections and categories in the evaluation checklist (exemplified through conceptual framework and discussed in chapter 4), and finally ‘who-where-and when’ are revealed through the empirical data of the formative review cycles. Another contribution of design-based research is what some researchers have called ‘ontological innovation’. MacKellar (2010) explains different researchers’ views on that issue. For example, diSessa & Cobb (2004: 84) proposed that “ontological innovations are attributions we make to the world that necessarily participate in our deepest explanatory frameworks”. Moreover, the ontological innovations that result from design-based research are explained by Gravemeijer & Cobb (2006: 23) to include the development of the conceptual framework which is also a result of this study:

The development of a conceptual framework to describe the phenomena under study is an essential part of a scientific endeavor. New categories, however, do not come readymade, and cannot simply be captured by writing down a definition. New categories have to be invented and embedded in a supporting theoretical framework. Defining scientific terms is more like finding and validating a new category of existence in the world.

All of these outcomes, are considered as theories, whether they are linked to setting understanding, the instrument development or the design procedures and guidelines, as well as their use. Edelson (2002) also explains the main theories of design-based research as “domain theories, design frameworks and design methodologies”. The domain theories are the generalizations “of some portion of a problem analysis” with their “two types... context theories and outcomes theories.” On one hand, the context theory is realized through the contextual challenges that are discovered while studying the context, and on the other hand, the outcome theory is obtained through the problem analysis which will eventually lead to certain outcomes, whether they are positive or negative. The design framework is “a generalized design solution” in the form of guidelines about the intervention. And finally, the design methodology “is a general design procedure” that “provides guidelines for the process rather than the product” (Edelson 2002: 113-115). Confrey (2006, cited in MacKellar 2010: 139) suggests that the ultimate goal in refining designs and generating explanatory theories lies in their ability to guide practice and that the ultimate measure of validity lies in their usefulness:

One cannot prescribe practices, but one can guide practice by means of explanatory frameworks accompanied by data, evidence, and argument. An explanatory framework is: (1) at best a model of likely outcomes; (2) closely connected to its theories; (3) as robust as its links to evidence from multiple sources of interaction within ecologically authentic settings; (4)...as valid as it is useful to others who are familiar and experienced in similar contexts

The conceptual frameworks, the research phases, the data collection cycles, the instrument design, all lead to the professional development of the researchers and participants, a unique result of DBR studies. Reeves (2000, cited in Cotton et. al. 2009: 1365) indicates that design principles “are not the sole outcome of the development research process”. A fundamental tenet of this type of research is “the dedication to providing direct benefits to all stakeholders within the context of the research”. Furthermore, Sandoval (2014: 18-19) refers to the “most recent characterizations of design research” which suggest “that it is an approach with certain commitments: the production of innovative learning environments, knowledge about how such environments work in the settings for which they are designed, and...some more fundamental knowledge about learning or teaching”. Whether the study results were called theories (Edleson, 2002), ontological innovations (DiSessa & Cobb, 2004) or

design conjectures (Sandoval, 2004), the most important issue is their operational contributions. Thus the combination of academic research (through phases 1 & 3) and practical design (phase 3) goes beyond the narrow interpretivist view of educational issues and problems as it offers not only comprehensive understanding of the investigated topic, but also tangible manifestations of the educational concepts in forms of designed instruments, conceptual frameworks, guidelines, working mechanisms and conditions of success.

6.9 Design-Based Research in Educational Settings

Design-based research is not only an approach for studying complicated problems and situations, but it is also an empowerment tool for all people who are involved with it. For researchers, it can be considered a professional development tool that enables them to understand the setting where they are conducting their studies, besides the experience and knowledge they accomplish throughout the whole DBR phases. Design-based research also enables the researchers to be an insider of the project setting as well as an insider of the different types of research paradigms, methodologies and methods. It helps the researchers to construct an overview of their studies in particular and research in general. DBR offers researchers themselves a lifelong project to work on and improve, as well as a revealing experience of how interventions in educational research work and how they should be enhanced during the design-based research study and even after its completion. DBR creates a sense of good obsession among researchers towards their research projects improvements and development. Eventually, DBR makes the experience interesting, revealing and useful and leaves the researcher with a vision and a mission of what to do next. For stakeholders, it also helps them to understand their problems and discover how they can solve them theoretically and practically instead of accepting them and struggling with them on daily basis. Thus, design-based research may well be the main future vehicle of change in educational research with its real problems and various contributions. Also, practitioners' involvement will lead to professional development and understanding of their learning-teaching contexts. For educators, DBR studies will allow them to keep track and improve previous recommendations and build on them. The DBR results can benefit educational research and studies as they form a kind of continuum that can be used to know what has been studied and what needs further investigations. In other words, DBR helps to build theoretical and practical grounds for educational studies where researchers and practitioners can obtain a chronological order of the 'applicable' theories

/models and assumptions; a quality that other types of educational research have failed to provide for the past century, as researchers tend to lose the essence of the accumulative aspect of knowledge that is available in scientific research, resulting in a sort of chaos where everything could be correct and accepted. Furthermore, DBR will help to transfer educational research studies from the 'implications' of qualitative studies to their actual applications and disseminations. Also DBR can help to lessen the confusion existing in education by presenting practical knowledge to educators, which is based on real experiences and familiar terms, concepts, interventions and solutions. Definitely, DBR can help to transform the design of the intervention from a mere attempt of the individual's concepts and views backed by some expert or research guidelines into a wider perspective and perception that incorporates the experiences of the users of the intervention and consideration of their contextual issues and problems. Some of these reasons are summarized by Walker (2006: 8) in his explanation of the main reasons for the origination of DBR:

We have seen no intellectual breakthroughs in research in education comparable to advances in medicine, engineering, and the sciences; nor have we seen any measurable improvement in teaching practices or student learning on a large scale...The second reason why some researchers and policymakers have begun to find design research attractive is the availability of promising new theories of learning and technologies through which these theories can be applied. Cognitive science, activity theory (or social constructionism), and brain research offer new perspectives on learning that may well be more powerful than the theories that have guided traditional research such as behaviorism...and conventional social psychology.

Barab & Squire (2004: 3-4) identify "seven major differences between traditional psychological methods and the design-experiment methodology" as following:

Design-based research focuses on understanding the messiness of real-world practice, with context being a core part of the story and not an extraneous variable to be trivialized. Further, design-based research involves flexible design revision, multiple dependent variables, and capturing social interaction. In addition, participants are not "subjects" assigned to treatments but instead are treated as co-participants in both the design and even the analysis. Last, given the focus on characterizing situations (as opposed to controlling variables), the focus of design-based research may be on developing a profile or theory that characterizes the design in practice (as opposed to simply testing hypotheses).

Despite its importance and compared to other research approaches, educational design-based research (DBR) is not discussed in educational research methods references due to the fact that it is relatively new (Knowlton, 2007: 209). For example, Creswell (2007) distinguishes between five qualitative approaches, but these do not include DBR (see also Denscombe, 2007). In consequence, PhD candidates usually discover this methodology through their supervisors or personal investigation and interest. Despite that, Bakker & Van Eerde (2015) state that “DBR is worth knowing about, especially for students who will become teachers or researchers in education” as “design-based research is claimed to have the potential to bridge the gap between educational practice and theory, because it aims both at developing theories about domain-specific learning and the means that are designed to support that learning” (Bakker & Van Eerde 2015: 2). Moreover, DBR is recommended as a suitable methodology for addressing many complex educational problems that should be dealt with in a holistic way (Plomp & Nieveen, 2007) such as the development of the teaching materials evaluation checklist in this study. And despite the importance of the “current international publication culture in the field of education” which “powerfully privileges descriptive knowledge” and which is “by all means extremely useful”, the “explanatory, predictive and normative theories are also needed to enable productive change” as the “current publication culture is insufficiently aligned to the knowledge needs in educational practice” McKenney & Reeves (2014: 153).

Bassey (1995), discussing the influence of other research, philosophies and theories on educational research, explains that “research in educational settings which aims to develop sociological theory, psychological theory, philosophical constructs or historical ideas is not educational research, but sociological, psychological, philosophical or historical research in educational settings” (Bassey 1995, cited in Hammersley 2007: 145). He advised other educational researchers “to leave parental home (if sociology and psychology were the parents) and stand firmly on our own ground” where the ground he refers to “is the educational process of the making of decisions and judgments by practitioners and policy-makers, from the standpoint of trying to improve” their educational settings and solve their own problems. In fact we can conclude that the advent of the post method era, critical pedagogy, autonomous learning, and reflective teaching are all leading to the search for better research methods that are compatible with all of these educational changes and challenges. Design-based research is one of the natural

evolutions and advances mentioned above in all fields including education. Hence, design-based research is a methodology with “enormous promise that may strike an optimal balance between the rigor educational researchers seek and the relevance that researchers and practitioners alike deserve” (Reeves 2011: 20).

6.10 Design-Based Research Limitations

As design-based research is still in its development stage, researchers may face some issues when using it as theoretical framework for their studies. Conducting more design-based research studies can help to solve some of the issues related to design-based research complexities. As discussed by Matthew et. al., (2014), some of these concerns include “uncertainty about the DBR process, uncertainty about how DBR differs from other forms of research” as “DBR is typically imagined as a form of qualitative research”. Other issues include “uncertainty about how DBR differs from design, or why design is not research while “DBR proponents seek to establish DBR as a distinct and valid form of research,” and finally “uncertainty about what might make DBR effective” (Matthew et. al.2014: 1-2). Some of these issues are also discussed by Van den Akker (1999:11-12) and include: the “tension in role division between development and research, isolating 'critical' variables versus comprehensive and complex design and generalization of findings”. These problems can be overcome, according to Van den Akker, through the research processes. The researcher has to focus on the developing phase at the beginning of the study and later shift the attention to the design. For solving the second problem he suggests “to adapt research foci or procedures in the methodology approach” and include “a careful description of both the evaluation procedures as well as the implementation context”. The role of the researcher and other problems are also discussed by Plomp (2007: 42), in the introduction on design-based research, to include the “multiple roles of the researchers” as “the researcher is designer and often also evaluator and implementer” and “complications of real-world setting” and adaptability”. He then offers solutions for each problem. For example, the multiple role of the researchers can be resolved through the use of outsiders as evaluators (peer and experts), and thinking critically throughout the whole inquiry through a robust research design. To deal with the complications of context, the researcher has to develop good relationships with different stakeholders and at the same time make use of his outsider position to be objective in the study, which will encourage the participants to cooperate. The adaptability is required for the design of the research project and the role of the researcher, and in both instances, the researcher has

to be prepared for the iterative cycles and the feedback from all the participants during the various iteration of the data collection. Some further limitations are mentioned by Van den Akker (1999: 12) in the form of questions that need researchers' attention when using this approach:

What does (rapid/evolutionary) prototyping imply for efficiency of the development process? Will it affect the balance between creative and systematic features of the approach? Does it reduce the relevance of preliminary investigations? To what extent does it influence the relationship between methodology (as prescribed in literature) and actual design activities in professional practices (can 'theory' keep up with 'practice', or will the gap even widen)?

Another problem mentioned earlier, which has been discussed by many researchers is the methodology of DBR, as Sandoval (2014: 18) demonstrates: “despite this boom in writing and move into the mainstream, there remains confusion about design research as a methodology.” In spite of that, some of the concerns raised in the above questions can be solved through a detailed plan of the research project and through controlling the different trajectories and phases as well as transforming the research process into a visual representation using charts and diagrams such as conceptual frameworks and concept maps that describe all the procedures and processes required throughout the whole research project. Within this study, it is noticed that some worries about the multiple roles of the researchers can be considered a privilege rather than a drawback. As Joseph (2004: 236) points out, a DBR context “creates a unique opportunity” for the researcher to observe “the ways that design questions, research questions, and questions of practice can feed and flow into one another.” So it is perhaps time to move off all of these debates about the “uncertainties” and to focus on conducting more DBR studies about various educational problems in different settings. Also, more publications on how to conduct DBR, its terminology and definitions are needed to encourage researchers to use it in their studies. As explained by Reigeluth & Carr-Chellman (2009: 5) the “use of the same term to refer to different things and different terms to refer to the same thing ...is confusing for all...from beginning graduate students to expert designers and researchers” so the availability of summaries or a dictionary of different terminology of DBR will remove such confusion. Also, the available sources that introduce design-based research for PhD students are insufficient. There is only one book that deals with the DBR phases, which is McKenney & Reeves (2012). The other sources are usually committed to certain

aspects or generalities of DBR, such as Van den Akker (1999), Anderson (2002), Herrington et.al. (2007), Richey & Klein (2007), Anderson & Shattuck (2012), Kennedy-Clark (2013), Easterday et. al. (2014) and Kelly et al. (2014). Thus, more inclusive and detailed references are needed to encourage students to utilize this methodology in their research projects. For example, the availability of an edited book that contains several parts on how to conduct a DBR study will help to save students time and ensure more reliable studies. The themes suggested include, for instance, sections or chapters on the historical background of design-based research, research problems, developing conceptual frameworks and related research questions, paradigm matters with explanations of their ontologies, epistemologies, methodologies, methods and instruments with clear explanations of educational research types, similarities and differences. Also, detailed descriptions of DBR characteristics and its different phases and their boundaries, as well as interpreting the results, the discussions and the concluding reflections, may encourage many researchers to use design-based research in their educational studies.

6.11 The Study Limitations and Future Dissemination

This study aimed to find out how a teaching materials evaluation checklist could be developed, the possible sources, the design guidelines and the validation method to review the designed checklist. Roberts (1996 cited in Harwood 2010: 381) suggested that materials evaluation checklists “should be regarded as illustrative and suggestive only” because of their lack of generalizability to other contexts. What made the previous checklists ‘context-specific’ is the absence of a solid framework which can be used as a unified starting point for the teaching materials checklists developers. Hence, designers depend on their own background and their knowledge about their context. Also, the methodologies used in developing these teaching materials evaluation checklists do not help the developers to test, review and validate their checklists with both experts and potential users at the same time. In fact, “improvements in design and analysis” of evaluation studies “can have only marginal impact on the integrity of evaluation studies because the inferences about the impact of interventions remain fundamentally dependent on the quality of the data collection” (Burstein et. al. 1985: 66). Thus, the methods of iterative data collection, where the product is reviewed by experts and users, are very important to ensure the validity and reliability of the developed teaching materials evaluation checklists. And because almost “all data collection defects are problems of

design” a method which combines research and design aspects is the appropriate one, as it minimizes the evaluation “faults that occur because of inconsistent data collection procedures, because of reactions to the data collection procedures, and because of insufficient oversight and quality control during data collection.” Solutions to all of these problems can be offered through formative review (a method used in design-based research studies) which provides detailed “final reports” which can also include “appendices containing the information” (Ibid) about all the cycles and procedures of data collection and analysis. It is hoped that this study will help to shed some light to the idea of basing the development of future teaching materials checklists on mutual and general sources (within the validated conceptual framework), even when the content and the items of the developed checklists are going to be different.

Despite its important contribution to the field of design based research, teaching materials evaluation in particular and educational evaluation in general, the generalizability of this study can only be achieved when a final summative review is conducted with larger numbers of users. The developed conceptual framework will need to be used by other teaching materials checklist developers to see if both experts and teachers can employ it in developing their evaluation checklists in different settings. These limitations can be resolved through repeating the formative review cycles using the same data collection instruments in different settings and with different users, and this option is provided within the formative review method. Researchers can repeat any cycle whenever they have time without having to start from scratch every time they need to develop or revise the developed checklist.

During this study, some issues about design-based research methodology were encountered. For example, with DBR, researchers cannot determine from the beginning if quantitative/ qualitative /mixed methods will be suitable, as each cycle may require one type of data or both. To solve that problem, a method that helps to answer the research questions can be selected, such as formative review that was used in this study. Also, the results from some data collection analysis may require using new methods and may lead to others. Likewise, the interval between one cycle and another to make the appropriate changes on the designed prototype before embarking on a new cycle can delay the research project processes. Another issue is related to intersection of the phases. For example, the analysis and exploration phase may include a construction and design phase. Also, the evaluation and reflection phase may need an analysis and exploration phase at

the beginning. Some of the terms used in design-based research, such as 'intervention' can be alleviated and probably replaced by a general term that suggests novelty, like 'innovation', to avoid confusion of design-based research studies with other educational research. Though design based research can make use of quantitative, qualitative and mixed methods as well as instructional and design models in collecting data, evaluation methods as formative review can be the most appropriate one. This method (e.g. evaluation) can help researchers solve two of the DBR dilemmas as (how to plan in advance for data collection within the DBR different phases and when to stop iterations). The formative review with its clear four cycles can guide researchers in such a confusing situation. They will be able to decide from the beginning that they will have to plan for four cycles (expert appraisal, one-to-one, small group & field testing) which will simplify the study processes and procedures.

6.12 Final thoughts

Conducting design-based research for the first time may be difficult, but, certainly the person who started the study will be different by the end of his/her project in terms of knowledge about the topic, conducting different types of research and understanding educational settings and their problems. Regarding the topic, researchers will have a full understanding of all the aspects related to it. Also, researchers will gain knowledge about different research types, as they will have exposure to different research paradigms, epistemologies and ontologies as well as to the three main types of research methods, qualitative quantitative and mixed methods, and their different techniques of data collection and analysis. In educational contexts, the main purpose of research studies will change from filling a gap in the literature into filling a gap in the real setting through focusing on the educational problems and defining their practical solutions. The researchers will be familiar with the details of the context and overview of the whole setting. Van den Akker et. al. (2006: 17) indicate that “the underlying philosophy of design research” is about understanding the badge “if you want to change something, you have to understand it, and if you want to understand something you have to change it.” There can be no useful and lasting purpose or aim if only one person (i.e. the researcher) can understand the problem and strive for changing it. Thus, the pursue of change and reform of educational systems and contexts for the better will take longer periods of time and endless efforts unless the stakeholders, the surrounding environment and the authorities are understanding their problems and are willing to participate and support

such transformations. Certainly, solving problems and improving practices in the educational settings will require the efforts of more than one person, or a group of researchers; they will require collaboration and cooperation between all stakeholders, which is provided through using pragmatic methodologies like DBR. In this study, the participation of stakeholders was through basing the checklist items on setting needs, on their wants and through allowing them to review and validate the checklist along with the experts. Unlike some other research studies, the final products, programmes or processes have to be spread, disseminated and incorporated after the end of the research projects. It is expected also that the products and innovations resulted from design-based research studies will encourage practices like continuous assessment, reviewing and quality management procedures in the educational environments and institutions through all the involved parties.

To conclude, it can be assured that “workplace conditions and practitioner problems are not likely to disappear and those that do exist will continue to fuel new research efforts” (Richey & Klein 2007: 149). Accordingly, this study is just the beginning of a new visualization of teaching materials evaluation in English Foundation Programmes. A new conception that sees materials evaluation through the clear understanding of their sources, the design processes as well as involving the stakeholders which will eventually lead to the professional development of different practitioners. Using design-based research in developing the teaching materials evaluation checklist in this thesis involves planning for “three main stages of implementation” mentioned by McKenney and Reeves (2012: 160) which include: “adoption, enactment and sustained maintenance”. In consequence, the development processes of the teaching materials evaluation checklist and any designed innovation or instrument never finishes. Its completion is just the beginning for other dissemination and implementation procedures as the results of these studies are initiated to survive and evolve through time in form of tangible instruments, processes and heuristics that can be used by different users and researchers and which can form the basis for future research and investigations.

Appendices

Appendix A: Subject Matter Experts Short CVs

Prof. Brian Tomlinson: Starus: Honorary Visiting Professor, University of Liverpool, TESOL Professor, Anaheim University, and President of MATSDA. **Areas of Expertise:** Materials Development and Evaluation, Curriculum Development and Evaluation, the Roles of Inner Speech and Visual Imaging in the Learning and Use of Languages, Teaching Language through Literature, Reading Research and the Teaching of Reading, Teacher Development, Language Acquisition and Development...Other Recent Positions: Consultancies of 19 institutions worldwide including: Consultant for Sultan Qaboos University Textbook Evaluation Project. Research Examining and Supervision 2002-2015 over 24 PhD studies+ other advising posts. **Recent Employment History:** 2015 -Honorary Visiting Professor, University of Liverpool. 2012 -TESOL Professor, Anaheim University. 2010 - 2015 Academic Director TEFL International. 2010 - 2011: Visiting Professor, Azad University, Oxford. 2008 – 2015 Visiting Professor, Leeds Metropolitan University. 2007 Curriculum Specialist, Sultan Qaboos University, Muscat. 2000 – 2007 Reader in Language Learning and Teaching, Leeds Metropolitan University (Research Manager, Head of the Post-Graduate, Research and Consultancy Unit and supervisor of PhD theses and MA dissertations).1997-2000 Senior Fellow and Research Coordinator, Department of English Language and Literature, The National University of Singapore (Module Leader on MA and BA courses and supervisor of PhD and MA dissertations). **Recent Publications:** over 11 books and 31 chapters in academic books...21 articles in Internationally Refereed Journals and 13 articles in Locally Refereed Journals.

Dr. Saleh Al Busaidi: Associate professor, Department of curriculum and instruction College of Education, Sultan Qaboos University).PhD from the University of Illinois at Urbana-Champaign, USA, 2003, MA from Exeter University, UK, 1997. BA from Sultan Qaboos University, Oman, 1995 .**Specialization:** Curriculum studies. **Research Interest:** Teacher education, Curriculum design, Teaching English as a foreign language Learner autonomy. **Selected Publications:** Tuzlukova, V. & Al-Busaidi, S. (2016). Research on technology-based language education in the Sultanate of Oman: perspectives for student skills' enhancement. *Journal of Teaching English for Specific and Academic Purposes*. 4(1), 1-8....Al-Busaidi, S., Aldhafri, S. & Büyükyavuz, O. (2016). Effective

university instructors as perceived by Turkish and Omani university students. *SAGE Open*, July-September, 1 (8). DOI: 10.1177/2158244016662900.

Al-Busaidi, S. & Al-Saqqaf, A. (2015). English spelling errors made by Arabic-speaking students. *English Language Teaching*, 8 (7).....Al-Busaidi, S. & Sultana, T. (2015). Critical thinking through translated literature in the EFL Omani Class. *International Journal of English and Literature*, 6(1), 16 – 22....Al-Busaidi, S. & Al-Mamaari, F. (2014). Exploring university teachers' understanding of learner autonomy. *Theory and Practice in Language Studies*, 4(10), 2051 – 2060....Al-Busaidi, S. (2013). The integration myth: reading and writing. *Journal of Social Sciences & Humanities*, 21(3), 1147 – 1156....Borg, S., & **Al-Busaidi, S.** (2012). Teachers' beliefs and practices regarding learner autonomy. *ELT Journal*, 66(3), 283 – 292.

Jayakaran Mukundan, PhD: Appointed Professor, English Language Teaching, 2011. Winner, National Award for Academic Excellence, 2014. Recipient, Vice-Chancellor's Award, 2013. Awarded Anugerah Pengajaran Putra (Putra Teaching Award), 2007. 5 time winner of Excellent Service Award, Universiti Putra Malaysia (2014, 2010, 2006, 1999, and 1998). Involved in more than 28 funded research projects. Publications include more than 80 journal articles, 23 chapters in books, 5 books and 25 edited books. Supervised and helped successfully graduate more than 20 PhD students and 30 Masters students. Made more than 150 conference presentations. Inventor of the 1st computerized tool for evaluating ELT textbooks, 2010 and software for retrospective evaluation of textbooks; Gold Medals at the British Invention Show and IEANA (Germany). Inventor of 1st on-line textbook evaluation checklist – www.elt-tec.com; 2011. Research Awards; 5 international awards; 3 national; 29 PRPI (UPM). **Major International Linkages:** Visiting Professor, State Islamic University, Jakarta, Indonesia, 2016; Visiting Research Fellow, Leeds Metropolitan University (since 2005); Visiting Professor, Ho Chi Minh Open University (since 2011); Director, Extensive Reading Foundation (since 2000); Visiting Research Fellow, Monash University, Melbourne, Australia, 2008; Co-Chair, Regional Creative Writing Group (since 2003); Co Chair, The 1st Extensive Reading World Congress, Kyoto University, Japan, 2011. **Major National Linkages:** Visiting Professor, Management and Science University (MSU), since 2006; Expert on the Minister of Education's Advisory Panel, since 1989; Expert on the MARA Junior Science

Schools ELT Panel, since 2011; Evaluator on LAN (National Accreditation Board) panel, since 2000; Founding Chair of 5 major international conferences (IMELT, MICELT, ICELT, ELT Materials, Creative Writing); Founding Chair of 2 major international symposiums (Creative Writing and ELT Materials); Organized more than 15 international conferences as Chair; Successfully implemented School Adoption and international student mobility Projects.

Dr. Vahid Nimehchisalem: Vahid Nimehchisalem has been involved in English language teaching since 1994. He is a senior lecturer in the Faculty of Modern Languages and Communication, Universiti Putra Malaysia. His areas of research interest include assessing writing and English language teaching materials. Both research projects he is currently involved in are in the area of self-assessment in ESL writing. He is chief editor of the International Journal of Education and Literacy Studies, an editorial team member of the International Journal of Applied Linguistics and English Literature, and is a regular reviewer of articles submitted to *Pertanika JSSH* and a few other journals. **Editor-in-chief:** International Journal of Education and Literacy Studies.

Appendix B: Authorization Letter

التاريخ، 18 / 11 / 2013م

الرقم، 1/6 / 1304 / 2013 ص.خ

المحترمين

الافاضل/ عمداء كليات العلوم التطبيقية

السلام عليكم ورحمة الله وبركاته وبعد،،،

للموضوع، تسهيل مهمة الباحثة، منى كشوب

إشارة إلى إبتعات الفاضلة/ منى كشوب لنيل شهادة الدكتوراه في المملكة المتحدة، فإننا نرجو منكم التكرم بتسهيل مهمة المذكورة لتطبيق أدوات البحث المعنون بـ:

**Evaluating ELT Materials in the Colleges of Applied Sciences Foundation
Programme in Oman**

و تتضمن الدراسة توزيع استبانة و إجراء مقابلات و مشاهدات صافية. شاكرين لكم دوام تعاونكم

و تفضلوا بقبول فائق الإحترام و التقدير،،،



د. عبدالله بن علي الشبلي

مدير عام كليات العلوم التطبيقية



نحنو تعليم عال ذي جودة عالية يلبي متطلبات التنمية المستدامة

سلطنة عُمان ص.ب: ٨٢ روي - الرمز البريدي: ١١٢ - هاتف ٢٤٢٤٠٥٨٠ / فاكس ٢٤٢٤٠٥٧٨
Sultanate of Oman, P.O.Box: 82 Ruwi, PC 112, Tel:24340580 / Fax 24340578 www.moha.gov.om

Appendix C1: Experts Questionnaire for the conceptual framework

Dear Dr. / prof.,

This short survey is intended to check and validate the developed conceptual framework that will be used to guide the designing of the evaluation checklist for the teaching materials as well as the study's procedures and phases. Your expertise will enable the researcher to detect any design or theoretical flaws in the framework. I would appreciate if you could answer the open-ended survey questions within two to three weeks.

Based on the above conceptual framework:

1. What is your first impression in terms of the framework's practicality for target users (teachers, programme coordinators and experts)?
2. What do you think of the procedures of the framework development that led to the development of the 1st prototype of the teaching materials checklist?
3. What are the items that you think should be deleted or changed? Why?
4. What are the missing points or stages that you think should be included in the framework? Why do you think they are important?
5. What are the items or the processes in the framework that you think are not clear? What are your suggestions to improve them?

Many thanks for your appreciated and valuable participation and looking forward to your precious comments and recommendations.

Appendix C2: Experts feedback on the conceptual framework

Questions	Expert 4 Dr Tomlinson	Expert 2 Dr Saleh	Expert 3 Dr Vahid	Expert 1 Dr Mukundan
1. What is your first impression in regards to framework's practicality for target users (teachers, experts and other materials evaluators)?	<p>I think it's principled and coherent but it needs to be made clearer if it's to be used by other people. At the moment it's only really useful to you or to a reader of your thesis who's already been provided with detailed information about what you did. If you want to make it practical for others to use it I think you need to:</p> <p>Explain some of the terms (e.g. English language inventories; Items integration). Your framework provides a useful summary of your procedures but it would be of little help to someone unfamiliar with the terms you use.</p> <p>Provide a brief description of what teachers, experts and other materials evaluators would actually need to do at each stage. For example, which literature do they review and which inventories do they use; and how do they ascertain student needs or integrate items?</p> <p>Indicate how the Theoretical aspects and the Contextual aspects are ultimately integrated when actually constituting the Prototype Checklist. Will there be two checklists (i.e. A + B) or one?</p> <p>Indicate whether the ultimate Formative evaluation is of the materials or of the Prototype checklist.</p>	<p>I support the involvement of the various stake holders in the development of the checklist. As the researcher mentioned in her review of the literature, very few checklists were developed based on the perspectives of the stakeholders. It would of course be quite challenging to reconcile the different views!</p>	<p>When I first saw this framework, I asked myself, "why is it linear"? To my experience, instrument development is circular (allowing iterations), or to be more precise, heuristic and recursive.</p>	<p>The framework seems to look like it is comprehensive, bridging theoretical considerations to context issues. I do hope that there is trialling involved under "contextual aspects" (or perhaps you should create a third front for this?) so as to flesh out aspects that stakeholders feel are most important as criteria for evaluation. If trialling isn't accounted for, in that case the framework isn't very clear in this aspect. You need to describe what goes on in "Theoretical Aspects" and in "Contextual Aspects". You also need to clarify why these 2 aspects were considered, why not others. Justifications are necessary. There are others and why aren't they included? Consider reading my PhD (which is on the Composite Framework for Textbook Evaluation, Mukundan, 2004) and also articles on RETROTEXT-E – a computerized frame work for Retrospective Evaluation of textbooks.</p>

<p>2. What do you think of the procedures of the framework development that led to the design of the first prototype of the teaching materials checklist?</p>	<p>They constitute a coherent and progressive framework which evolves organically and progresses logically. I'm not sure though why you start with an evaluation checklist. How can a checklist be developed before the other procedures have been followed?</p> <p>There are some inconsistencies in the framework For example, Theoretical aspects is the heading for a procedure (i.e. reviewing the literature) whereas the equivalent heading on the right hand side (i.e. Contextual aspects) is a just a category heading.</p>	<p>The procedures look fine generally speaking. However, I still do not see the need for the development of two prototypes. Why not develop one from the beginning? Also, I am not sure if the researcher would find it easy to combine all the existing checklists into one single checklist!</p>	<p>The procedures sound comprehensive, but still open for discussion. One may question why should you create the two separate prototypes (A and B) rather than only one?</p>	<p>There needs to be clarification as to how items are sourced from all the components within the framework. How do you frame items? Are you going to sift through past instruments and model your items (from your own analysis from data) based on past ones? Or are you going to create your own through focus groups? The procedures look static in a Framework. The researcher has to illustrate the workings of the Framework in detail.</p>
<p>3. What are the missing points or phases that you think would need to be included in the framework? Please support your arguments.</p>	<p>There is no stage where the evaluators brainstorm their own beliefs about language acquisition or refer to their own experience of language learning and teaching. Without this stage the evaluators are dependent on the theories and dictates of academic researchers, of examiners and of curriculum developers and they are ignoring the invaluable resource of themselves.</p> <p>What is really missing is some indication of how all the information gathered can be combined when formulating evaluation criteria. How, for example, can the data on teachers' needs be combined with information about learning theories? When the evaluator actually writes the checklist what categories are used and what is prioritized?</p> <p>Establishing a set of criteria for evaluating the evaluation</p>	<p>I do not see a stage/section for refinement of the checklist based on the piloting/evaluation.</p>	<p>Since you are in fact developing a test, I would suggest following a more test design framework like Fulcher's (2010) 'Test design cycle' which starts by specifying your test purpose, followed by specifying the criteria, defining the construct, item writing, prototyping, field-testing, inferences and decisions.</p> <p>[See chapter 4 in Fulcher, G. (2010). Practical language testing.</p>	<p>Since you said in your proposal that trialling that is done in Oman is "tedious" why don't you then put aspects of what is found in trialling as the 3rd Source of item construction within your framework. I do believe that so far the "contextual aspects" covers interviews, etc. That 3rd phase will seriously look at the practical aspects of textbook use and your framework would have even better novelty value.</p>

	<p>instruments would be a very useful additional stage.</p> <p>How do the evaluators select from the mass of data what should inform their checklists?</p> <p>If everything is included the checklist will be so unwieldy as to be unusable?</p> <p>I think you need to indicate how the framework should be used.</p> <p>Should the evaluators work through all the stages on one side first and then work down the other side or should they work across. If the latter, does it matter if it's from left to right or from right to left? Numbering the recommended sequence would help.</p> <p>I think you also need to indicate that using this framework isn't just a straightforward progression from stage to stage.</p> <p>It's a recursive process involving going both forwards and backwards and making numerous revisions.</p>		<p>London: Hodder Education.]</p> <p>This framework is more manageable and will help you see your final phase 'evaluation' more explicitly and specifically.</p>	
<p>4.What are the items that you think should be deleted or changed? Why?</p>	<p>At the moment you are mixing analysis and evaluation by including, for example, both content (e.g. Language inventories) and pedagogical approach (e.g. Learning theories). I would have a framework for an analysis of the materials which would include those procedures which would facilitate the framing of criteria for materials analysis (i.e. questions with factual answers about what the materials contain and what they ask students to do). I would also have a framework for an evaluation of the materials which would include those procedures which would facilitate the framing of questions for materials evaluation (i.e. questions</p>	<p>None except for the development of two prototypes (see my answer to Question 2 above).</p>	<p>As suggested above, it sounds more manageable and less arduous to create the instrument first and then refine it based on the stakeholders' needs, interests, and/or views (rather than coming up with two instruments and then integrating them).</p>	<p>Several things are unclear at level 4 of the framework:</p> <ul style="list-style-type: none"> • Items integration (shouldn't it be item development). Also "Data Collection and Analysis" – what will the outcomes of this be. Shouldn't these two be about "sourcing" of items for your instrument? These lack in clarity. • Why call these Prototype A and Prototype B? There should only be ONE prototype. All the other phases will be developmental aspects that lead to the Prototype.

	<p>inviting assessment of the likely degree of effectiveness of the materials). An example of an analysis question would be, 'Do the materials include information about the functions of the passive voice?' An example of an evaluation question would be, 'To what extent are the reading texts likely to stimulate affective engagement?' Evaluation question can be scored from 1-5; analysis questions cannot. Mixing them up can cause great confusion.</p> <p><u>I would delete:</u></p> <ul style="list-style-type: none"> -The first evaluation checklist (I don't see any point in it) -All the grey boxes (they are used inconsistently, sometimes to specify, sometimes to clarify and sometimes to exemplify) <p><u>I would change:</u></p> <ul style="list-style-type: none"> -Theoretical aspects to Literature Review -Contextual aspects to Needs analysis and delete Stakeholders' needs -Language theories to SLA principles -Learning theories to Learning and Teaching principles -English language inventories to Curriculum and examination specifications -Setting needs to Institutional needs -Data collection and analysis to Data collation and analysis <p>Formative evaluation to Evaluation of the materials or Evaluation of the instruments of evaluation (whichever is intended)</p>			
--	--	--	--	--

	I think making these deletions and changes would make the framework much clearer and much more useable for evaluators.			
5. What are the items or the processes in the framework that you think are not clear? What are your suggestions for improving them?	I think I've specified all of these above.	There should be lines connecting the blue boxes to all three boxes above them to show that each of the prototypes is based on the three sources. The box titled "Setting needs" is not clear.	The final phase, 'evaluation' needs to be expanded. How is this evaluation going to be and how is it going to be formative?	Perhaps Language Theories and Language Learning Theories can be combined otherwise you may be stretching this too far. Also, it is very unclear what is meant by "English Language Inventories" and what is the purpose of having it in the framework. I do notice that the justifications for the inclusion of all these components in the Framework are not explained very clearly.

Appendix D1: Brain writing Sessions Sheet (Teachers)

<p>What are the characteristics of a "good" English language textbook?</p> <p>What are the activities and tasks that you would like to be included in your English language textbook?</p> <p>How do you usually judge the effectiveness of English language textbooks while using them?</p> <p>Which content and skills would you like to see more in English language textbooks?</p>				
Participants	Round	Idea 1	Idea 2	Idea 3
A	1			
B	2			
C	3			
D	4			
E	5			
F	6			

Appendix D2: Brainwriting Sessions Sheet (Students)

What are your goals/ purposes for studying English? ما هي أهدافك من دراسة اللغة الإنجليزية؟				
What are the language items and skills that you would like to see in your textbook? ماهي المهارات اللغوية التي تتمنى أن تراها موجودة في كتاب اللغة الإنجليزية؟				
How do you usually study for your English Course? كيف تذاكر مادة اللغة الانجليزية؟				
What are the strategies and techniques that you use to succeed in English learning? ماهي الطرق والوسائل التي تستخدمها للنجاح في تعلم اللغة الإنجليزية؟				
Participants	Round	Idea 1	Idea 2	Idea 3
A	1			
B	2			
C	3			
D	4			
E	5			
F	6			

Appendix D3: Instructions for the brain writing sessions

At the beginning of the session:

1. Preparing the setting (the room, the tables, the brainwriting forms and pens
2. Explaining what is Brainwriting compared to brainstorming which is familiar to the participants.
3. Explaining the 6-3-5 brainwriting method.
4. Introducing the topic to the participants (teaching materials in the English language programme) and the four questions that they will generate ideas for.

The instructions for the participants:

- The facilitator (the researcher) is going to explain all the procedures to you as well as keeping the time for each round.
- Please make sure to write clearly complete sentences (3-6 words)
- Please avoid using vague language and abbreviations
- The first round will take five minutes in order to write 3 ideas
- Then the form is passed to the person next to you (on your right)

-Please pass the forms silently and do not speak or discuss ideas with the person next to you.

-When you receive the form from the other teammate, you can read their ideas and then write your own based on them or write new ones.

-When the six forms are filled up with ideas, the session will stop and data analysis will start immediately using cluster analysis technique.

Data Analysis:

1. Cluster technique:

- When the brainwriting session finishes, the participants will start grouping the ideas into clusters or categories.
- Each category is given a label and ideas are put under the related category.
- To reach a consensus about merging or eliminating some ideas, the affinity diagram will be used.

2. The affinity diagram process:

- The ideas resulted from the brainwriting session will be placed on a large table.
- The participants will put the similar ideas or the “affinity groups” together.
- Participants “can move any item” but they should keep silence “while grouping items” and “do not discard duplicates”
- The participants will then label each affinity group.

Finally, the ideas will be prioritized by the researcher in order to be used in developing Prototype B which will be (along with prototype A) the basis for the teaching materials evaluation checklist.

Appendix D4: Pictures from the brainwriting sessions



Appendix D5: Themes of the brainwriting sessions

➤ Students (themes in form of questions)

- Do the materials focus on all the four learning skills?
- Are the sub-skills (micro skills) incorporated with the appropriate main skill (macro skills)?
- Does the materials content include all the required items of the language?
- Does the materials incorporate the appropriate strategies and techniques in delivering the language?
- Does the materials utilize the available sources like internet, mobile phones, newspapers and magazines?
- Are there sufficient activities of dictionary use, dictation, and spelling games?
- Are there enough illustrative drawings, pictures, maps and infographics?
- Are there entertaining activities like songs, short films and documentaries?
- Do the materials allocate enough time for students to speak in class and be involved in teaching and presenting parts of the lessons to the class?

➤ Teachers themes (brainwriting sessions)

Questions	Teachers recommendations
<p>a. To what extent do the teaching materials consider the general qualities and preferences recommended by teachers?</p>	<ol style="list-style-type: none"> 1. Enjoyable for students and teachers 2. Can be modified through the availability of soft version of materials 3. Up-to-date teaching methods that incorporate gamification of learning and enable learners' autonomy 4. Curriculum with lots of recycling/clear aims and Outcomes that align with it 5. Publisher available for questions and feedback 6. attractive design and layout and the font size is readable
<p>b. To what extent do the materials comply with teachers' recommended tasks and activities?</p>	<ol style="list-style-type: none"> 1-Pronunciation activities 2-More listening activities (3 minutes or longer) 3-Personalized speaking activities 4-Graded reading activities 5-Paraphrasing tasks 6. Peer , pair work, group work activities 7. Low level sentence structure activities 8. Research-based activities 9. Reading comprehension activities 10. Skimming and scanning activities 11. Vocabulary and dictionary work tasks 12. Error correction and error analysis activities

<p>c. To what extent do the teaching materials comply with teachers' quality criteria?</p>	<ol style="list-style-type: none"> 1. Organized in logical way 2. Easy and practical exercises for students and teachers 3. Variety of topics: Wide range of topics/ engaging topics and materials 4. Content is interesting 5. Culturally appropriate 6. Easy adaptable materials 7. Aligned to curriculum objectives and students' needs
<p>d. To what extent do teaching materials consider the teachers' preferred content and skills?</p>	<ol style="list-style-type: none"> 1. Content that exploits CLIT (content language-integrated learning) 2. Utilizing E-learning: List of online sites for further learning, Soft copy of vocabulary list 3. Reading skills as Skimming/ scanning 4. Vocabulary and Dictionary use skills 5. Organizing a paragraph skills as paraphrasing and summarizing 6. Different genres of writing (essays, letters, emails) 7. Research skills as analyzing graphs and tables 8. Research which lead to problem solving, problem solving and critical thinking 9. Topics that are relevant to students as modern lifestyle (technology/ life skills), 10. Topics that prepare them for their majors (vocabulary)

Appendix E: Short survey of the institutional needs

1. What is your impression of the core and supplementary course books currently being used amongst teachers and students at the Colleges of Applied Sciences, English language programme?
2. In your opinion, what methods can be utilized to evaluate the teaching materials (the English language textbooks) used in the Foundation Programme?
3. Please rank the following in terms of their importance when selecting or evaluating English language teaching materials in the Foundation.
Foundation Programmes National standards.
Students' Proficiency levels.
The price of the textbooks.
Students' majors and degree programs.

Appendix F: Formative review stages and instruments

Formative evaluation instruments	Participants	Formative evaluation Purpose	Formative evaluation Methods	Prototype
Developer screening	The researcher	Discovering obvious errors	Developer screening through using 3 checklists on how to develop evaluative checklists (Tomlinson, 2013 ; Wilson, 2013; Bichelmeyer, 2003 and Stufflebeam, 2000)	Version 1
Expert Evaluation	Experts (n=4)	Relevance (Content validity), Consistency (construct validity), General design, Format and Structure	4 open-ended survey questions	Version 2
One-to-one Evaluation	Teachers (n=3)	Clarity, Appeal, errors, Practicality and Usability	Users annotated analysis of the completed checklist with specific instructions	Version 2
Small group evaluation	Experts from Other institutions (n=2) Teachers (n=6)	Effectiveness, appeal, Usability and Practicality	Micro-evaluation in a small group through survey questions on 3 main issues: 1. Clarity 2. Practicality and 3. Usability.	Version 3
Field Test	Teachers (n=4) Foundation programme coordinators (n=6)	Effectiveness and acceptance by the users	Field test through using the finalized checklist in evaluating 2 proposed textbooks (one familiar textbook and other to the users)	Version 4(the final one)

Appendix G1: Experts feedback Questionnaire on the conceptual framework

Dear Expert,

This is the first draft (prototype) of the checklist for the evaluation of English language materials in the foundation programmes. Your feedback is of great importance in this initial stage of the checklist development as it will help in amending and improving its **content, format and structure** in order to make it more comprehensible and useable for its potential users.

Please answer the following questions with regard to the attached evaluation checklist:

1. What are your suggestions regarding the **content** of the checklist in terms of its:
 - c. Inclusiveness of all the necessary items for a teaching materials evaluation checklist.
 - d. The precision of its words and terms.
2. What are your suggestions regarding **structure** of the checklist in terms of:
 - c. The grouping and sequencing of the specific items within the main categories.
 - d. The transparency of the checklist's layout with reference to the main headings and sub-headings, numbering, organizations of items and attractiveness to its prospective users.
3. The reliability of the checklist in terms of its generalizability to other contexts and settings.
4. Further suggestions on:
 - c. Any other sources, categories or items to be added to the checklist?
 - d. Any other methods that can be used to judge or evaluate the checklist?

Appendix G2: Experts feedback on the Checklist prototype 1

Questions	Expert 1(Dr. Mukundan)	Expert 2 (Dr. Saleh)	Expert 3 (Dr. Vahid)	Expert 4 (Dr. Tomlinson)
<p>3. What are your suggestions regarding the content of the checklist in terms of its:</p> <p>2. Inclusiveness of all the necessary items for a “general English” teaching materials evaluation checklist.</p> <p>3. The precision of its words and terms</p>	<p>First need to specify why you need to label this “General English”?</p> <p>Mostly you have covered most criteria</p> <p>Generally, satisfactory</p>	<p>1.a It is not clear who this checklist is intended for. The checklist is immensely too long. Teachers will certainly find it time consuming to use! I wonder if a shorter, teacher friendly checklist could be developed based on it.</p> <p>1.b The wording is clear and precise. I like the idea of showing the theoretical basis for each of the criteria (what to look for). There were just a few cases that require reconsideration or rewording.</p>	<p>The checklist covers most of the important aspects of general English learning-teaching. However, it is difficult to state this confidently. I would suggest following a well-established taxonomy to make sure that no aspect is left out. When one looks at this checklist through Bachman and Palmer’s (1996) components of communicative competence, one sees that the checklist covers most (if not all) of the components.</p> <p>Some of the items could be reworded more effectively. Some items look more like notes than clear-cut, unidirectional and terse items. I have left comments in the other file showing some of these items that I’m concerned about.</p>	<p>a. It’s a comprehensive list. I’d be tempted to delete some of the less useful criteria to make the checklist more user friendly.</p> <p>b. Sometimes you use the words of experts who haven’t used them with precision or clarity. I’d suggest using your own words and being a little more constructively critical of the experts.</p>

<p>4. What are your suggestions regarding structure of the checklist in terms of:</p> <p>e. The grouping and sequencing of the specific items within the main categories.</p> <p>f. The transparency (<i>what do you mean?</i>) of the checklist's layout with reference to the main headings and sub-headings, numbering, organizations of items and attractiveness to its prospective users.</p>	<p>I do not see the Relevance of (3)-Institutional needs</p> <p>We can only evaluate on this once the proto-type is ready</p>	<p>See comments on checklist</p>	<p>The sections and sub-sections look inclusive but since they come from different sources, the developer should make sure they do not overlap. I have left many comments related to this question in the other file.</p>	<p>The structure and sequencing is logical and clear.</p>
<p>5. The reliability of the checklist in terms of its generalizability to other contexts and how it can be adjusted to suit various settings</p>	<p>Reliability can only be established on testing the instrument</p>	<p>The checklist is too long. This might partly because we do not know the ultimate user of the checklist. Who is it for exactly? If teachers are going to use it, then it is way too long and complex. You need a much simpler and more teacher/user friendly form</p>	<p>There are a few items that could be worded more clearly and explicitly. They may lower the reliability of the instrument. What should be done at this point to increase the reliability of the instrument is avoiding items that look into more than one particular subject matter. Such items will confuse the user who will not know what exactly to consider when marking 1, 2, or 3 in this case.</p>	<p>I'd divide the criteria into universal criteria and local criteria. The universal criteria apply to any learner anywhere and are therefore generalizable to other contexts. The local criteria are specific to the particular context of your evaluation and are not transferable to other contexts without modification. Universal criteria derive from principles of language learning. Local criteria derive from a profile of the particular learning context.</p>

<p>6. Further suggestions on:</p> <p>e. Any other sources, categories or items to be added to the checklist?</p> <p>f. Any other methods that can be used to judge or evaluate the checklist? (<i>evaluate the checklist or the book?</i>)</p>	<p>Generally good. A comparison of other instruments can be done in the literature review</p>	<p>See comments on checklist</p> <p>See comments on checklist</p>	<p>I can't think of any at the moment by I mentioned one above and another in the annotated file attached.</p>	<p>Get a teacher to use it and then report their feedback to you on any problems they had in using it.</p> <p>Use it yourself to evaluate a specific textbook and note any problems that you have.</p> <p>General Comments</p> <p>The Impressionistic evaluation checklist is a mixture of analysis (i.e. 1 and 2) and evaluation (i.e. 3 and 4). Also I don't understand how to use the three availability columns in answer to suitability questions.</p> <p>Your checklist would be much more user friendly and reliable if you separated analysis from evaluation, if you used criteria to phrase questions, if your analysis questions were Yes/No questions and if your evaluation questions were answered on a scale from 1-5.</p> <p>e.g.</p> <p>Analysis</p> <p>1 Are all the components available?</p> <p>Evaluation</p> <p>1 To what extent are the topics likely to appeal to the learners?</p> <p>The close evaluation checklist would also be more user friendly and reliable if you used criteria to frame evaluation questions and if your evaluation</p>
---	---	---	--	---

				<p>questions were answered on a scale from 1-5. e.g. To what extent are the materials likely to: 1 help the learners develop confidence? 2 facilitate learner self-investment?</p> <p>Remember that with an analysis you are finding out what the materials consist of and what they ask the learners to do. With an evaluation you are predicting the likely effects of the materials on their users.</p>
--	--	--	--	--

Appendix H1: One -to-one review protocol

I. Preparations before administering the checklist prototype		
Topic	Items	Wording/ instructions for the users
A. Greeting and introduction	<ol style="list-style-type: none"> 1. Welcome the user 2. Introduce myself 3. Describe the aims of the study and the evaluation checklist being tested 4. Explain the goals of the session 5. Introduce the notion of paper prototyping 	<p>"Thank you for Participating. I'm Muna Kashoob. I am a PhD student at the University of York.</p> <p>The aim of this research is to design an evaluation checklist that will be based on an extensive study of literature in the area as well as the stakeholders' needs. The results of both literature review and needs analysis are used in designing and developing the English teaching materials evaluation checklist. Now, the first prototype of the checklist will go through 5 cycles or types of formative evaluation: developer screening, expert evaluation, one-to-one evaluation, small group evaluation and finally field testing. Your sessions are part of on-to one evaluation. I have already the evaluation package including the checklist. The purpose of today's session is for you to help me figure out how to make the checklist more user-friendly before we finish developing it.</p> <p>We will use the prototype of the checklist to identify problems in the checklist's "clarity, ease of use, sequencing and completeness" as well as its usefulness to you. These sessions will help also to determine revisions that can be made to improve it.</p> <p>A prototype is "an early sample, model, or release of a product built to test a concept or process or to act as a thing to be replicated or learned from. It is a term used in a variety of contexts, including semantics, design, electronics, and software programming". KREINTER, R., & Kinicki, A. (2007).</p>
B. Subject's role (Teachers)	<ol style="list-style-type: none"> 1. Explain what's expected of the user. 	<p>"The main objective is to discover weaknesses that can be amended to improve effectiveness" through the researcher's questions as well as your:</p>

		<p>a. spontaneous comments (where researcher records and also probes with follow up questions)</p> <p>b. Debriefing questions at the end of the session that can be used to ask about general strengths and weaknesses of the checklist.</p> <p>c. Observation of how you approach the checklist.</p>
C. Social concerns	<ol style="list-style-type: none"> 1. Explain that you're testing the checklist not them. 2. Reassure users about what will happen if they encounter any difficulties. 3. Reiterate how valuable this is and how much you appreciate their help. 	<p>Please “keep in mind that we're testing the checklist—we're not testing you—so if you run into any problems it's not your fault and it means that there's something we need to change”. I'll be sitting next to you, and I can help you if you want. We will have 3 sessions with three teachers and each one will help to discover more issues to amend or change. “We really appreciate having you come and help us out.”</p>
D. Set expectations	<ol style="list-style-type: none"> 1. Acknowledge the unfinished nature of the prototype that means we can make it better through users' feedback. 2. Explain that the design will evolve. 3. Explain that you will record their suggestions and will benefit from them in improving the checklist. 	<p>“The prototype still has some rough edges—we're still thinking through how it should work and some parts of it are incomplete. Before we cast it in concrete, we want to get some feedback about how well this design works... If you have suggestions we'll make note of them... When we get done with this series of sessions, we'll review everyone's feedback to help determine our priorities for the next” prototype.</p>
II. Introducing the evaluation of the checklist/ starting the evaluation		
Topic	Item	Wording/ instructions for users
Introductions	<p>Give the user the booklet that include the checklist prototype along with pens, pencils and extra sheets.</p> <p>Have users answer 2–3 questions about their background.</p>	<p>Here is your booklet, if you think you will need anything else please tell me now.</p> <p>First we will spend 2 to 3 minutes to let talk about yourself and your background, so please tell me about your experience and interests in teaching.</p>
Paper prototype orientation	<p>Explain what they're looking at.</p>	<p>“As I mentioned, here's the paper prototype” of the checklist you'll be working with. Please have a look at it first and then answer the questions about it.</p>

Interactions with the prototype	<p>Introduce the prototype.</p> <p>Explain how to interact with the prototype</p>	<p>"There are different ways to use this. Please tell us what makes sense to you, what's confusing, and any questions that come to mind. Your questions are especially valuable, but I may not answer them right away because our goal is to change the checklist so it answers them."</p>
They are in charge	<p>Remind the user that you're testing the checklist.</p> <p>Confirm ending time and that they can stop or take a break at any time.</p>	<p>"Remember that we're testing the checklist—we're not testing you. We'll end promptly" as soon as you finish, "but if you need to stop or take a break before then, just let me know. Are you ready to start?"</p>
Begin first task	<p>Hand users the first task.</p> <p>Clarify the task if it's confusing.</p> <p>If necessary, prompt the users to begin interacting with the prototype.</p>	<p>"Okay, here's the prototype that we would like you to do. Take 5 minutes to read the checklist and you can ask any question about it.</p> <p>As you read the checklist, please provide your feedback through your own comments and answering the researcher's questions</p>

Appendix H2: One -to-one data recording log

Data recording log

Teacher:

Personal information/ Background:

A. Subject' (Teacher's) comments	Page and category or item No.	Notes
1. 2. 3. 4.		
<p>B. Researcher's questions Clarity:</p> <ol style="list-style-type: none"> 1. Are there any categories, items, words or sentences that seemed confusing or ambiguous? Please underline vague items and rewrite them using other alternatives. 2. Is there any terminology that is puzzling to you? 3. Is there any item that is out of place? Please refer to the items that you think should be moved and locate their new place in the checklist. 4. Cross out the items that you think are repeated or unnecessary. 5. Are the checklist directions of use clear? Can you understand what is required from you? <p>C. Checklist completeness:</p> <ol style="list-style-type: none"> 1. Should more explanations be added? Which explanations can be added to improve the checklist? Where are the categories and items that need more explanation? (Suggest any missing items that you think are supposed to be included in the checklist.) 2. Do you feel that you need more help in understanding the checklist? <p>D. Difficulty /ease of use</p> <ol style="list-style-type: none"> 1. Can you use the checklist within a reasonable amount of time? 2. Do you think that you could use this checklist without help? 3. Do you feel that its use is challenging for ordinary teachers? 4. Where there any difficult parts? 5. Do you feel bored? Why? <p>E. Language obvious errors:</p> <ol style="list-style-type: none"> 1. Please refer to any obvious errors such as misspelling, incomplete sentences, poor grammar or any other mistakes that you could notice. 	<p>Teacher's answers</p> <p>Clarity:</p> <ol style="list-style-type: none"> 1. 2. 3. 4. 5. <p>Completeness:</p> <ol style="list-style-type: none"> 1. 2. <p>Easy or difficult:</p> <ol style="list-style-type: none"> 1. 2. 3. 4. 5. <p>Language problems:</p>	Notes
<p>F. Aspects /behaviors to observe and record</p> <ol style="list-style-type: none"> 1. Puzzled looks 2. Long pauses 3. Misunderstanding of some items, words or sentences 	Observations on specific items/parts of the checklist	Notes

	1. 2. 3.	
G. Unanticipated outcomes and comments	1. 2. 3.	
H. Debriefing questions	Answers	Notes
	1. What is your general comments and recommendations of the checklist that you think will make it more practical and effective? 2. Do you think that other teachers will be interested in this checklist? Why/why not? 3. Would you use this checklist to select or evaluate your textbooks? 4. What would you change in this checklist to make it better? 5. What did you learn from using this checklist?	1. 2. 3. 4. 5.

The protocol is based on:

1. Snyder, C. (2003). *Paper prototyping: The fast and easy way to design and refine user interfaces*. Morgan Kaufmann.
2. Beyer, B. K. (1995). How to conduct a formative evaluation. *Alexandria, VA: Association for Supervision and Curriculum Development*.

The Prototype definition is from:

KREINTER, R., & Kinicki, A. (2007). Organizational behavior: key concepts, skills and best practices.

Appendix H3: One -to-one Feedback (Teacher 1)

Teacher 1: Male, post graduate education, 16 years' experience, British

Tuesday, 8/ 11/ 2016 from 12: 20 till 1:05

	Page and category or item No.	Notes
<p>I. Subject's (Teacher's) comments</p> <ol style="list-style-type: none"> 1. Add feedback section at the end of the checklist so the evaluator can tell what is missing and get feedback from different evaluators on each topic of the checklist. 2. "objectives are clear" 3. "Grading level" and why you give that score to get "the teachers' thoughts" 	<p>At the end of the checklist or the last column (the evaluator can express his/her opinion through the added section.</p> <p>Language is clear</p>	<p>Seems that the teacher has a reasonable background that enabled him to understand the checklist</p>
<p>J. Researcher's questions (throughout the session)</p> <p>Clarity:</p> <ol style="list-style-type: none"> 6. Are there any categories, items, words or sentences that seemed confusing or ambiguous? Please underline vague items and rewrite them using other alternatives. 7. Is there any terminology that is puzzling to you? 8. Is there any item that is out of place? Please refer to the items that you think should be moved and locate their new place in the checklist. 9. Cross out the items that you think are repeated or unnecessary. 	<p>Teacher's answers</p> <p>Clarity:</p> <ol style="list-style-type: none"> 1. Easy to understand, but P.5 No. f: how to judge if this is available in the materials? 3. Brainwriting! 4. Nothing 5. P. 4 No.i: simple lives meaning differ in developing countries and rich ones. (Farmers in UK are rich whereas in developing countries are usually poor) Also travelers has different connotation in western countries (UK, for example, connected to gypsies) that differs from countries where English in a second language. 6. Yes 	<p>Notes</p>

<p>10. Are the checklist directions of use clear? Can you understand what is required from you?</p> <p>K. Checklist completeness:</p> <p>3. Should more explanations be added? Which explanations can be added to improve the checklist? Where are the categories and items that need more explanation? (Suggest any missing items that you think are supposed to be included in the checklist.)</p> <p>4. Do you feel that you need more help in understanding the checklist?</p> <p>L. Difficulty /ease of use</p> <p>6. Can you use the checklist within a reasonable amount of time?</p> <p>7. Do you think that you could use this checklist without help?</p> <p>8. Do you feel that its use is challenging for ordinary teachers?</p> <p>9. Where there any difficult parts?</p> <p>10. Do you feel bored? Why?</p> <p>M. Language obvious errors:</p> <p>2. Please refer to any obvious errors such as misspelling, incomplete sentences, poor grammar or any other mistakes that you could notice.</p>	<p>Completeness:</p> <p>1. P.3 No. h: Be more specific of what you are looking for. Reality of language use varies from one person to another. (needs more explanation)</p> <p>P.4 N. b (More details)</p> <p>P.6 No. b: Include the goals of learning English</p> <p>2. Identify what you are looking for: listing what you are looking for in all sections.</p> <p>Easy or difficult:</p> <p>1. yes</p> <p>2. Yes</p> <p>3. Could be especially if the teacher major is different field, and has taken TESOL to be able to teach.</p> <p>4. No</p> <p>5. No</p> <p>Language problems:</p> <p>1. P.5 No. b Are instead of is!!!! Second opinion</p>	
<p>N. Behaviors to observe and record (Observations on specific items/parts of the checklist)</p> <p>4. Puzzled looks</p>		<p>Notes</p>

<p>5. Long pauses</p> <p>6. Misunderstanding of some items, words or sentences</p>	<p>1. It seems that the teacher is able to understand the checklist</p> <p>2. There were no long pauses</p> <p>3. No major misunderstanding probably due to the teacher's post graduate degree.</p>	
<p>O. Unanticipated outcomes and comments</p>	<p>1.None</p> <p>2. None</p> <p>3.None</p>	
<p>P. Debriefing questions (at the end)</p> <p>6. What is your general comments and recommendations of the checklist that you think will make it more practical and effective?</p> <p>7. Do you think that other teachers will be interested in this checklist? Why/why not?</p> <p>8. Would you use this checklist to select or evaluate your textbooks?</p> <p>9. What would you change in this checklist to make it better?</p> <p>10. What did you learn from using this checklist?</p>	<p>Answers</p> <p>1. No more than the previous ones</p> <p>2. Some would/ some wouldn't: it depends on time and payment for the teacher.</p> <p>3. Yes</p> <p>4. Only as the mentioned above</p> <p>5. P.3 No. f: Cultural awareness- (useful information) the more you use the language, the more you use it.</p>	<p>Notes</p>

Appendix H4: One -to-one Feedback (Teacher 2)

Teacher 2: Female, Bachelor of education, 21 years' experience, American

Wednesday, 9/ 11/ 2016 from 12: 06 till 12: 38

	Page and category or item No.	Notes
<p>Q. Subject's (Teacher's) comments</p> <ol style="list-style-type: none"> 4. "Pretty extensive, quite thorough" 5. Activities that challenge learners to think creatively 6. Very broad, materials that suits all learners do not exist 	<ol style="list-style-type: none"> 1. Inclusive checklist 2. P. 3 No:g was admired by the teacher 3. P. 4 No. i: eliminate the word "all" 	<p>The teacher's background and long experience in teaching seemed to help understand the evaluation checklist. The more the teacher is experienced, the more he/she spends less time on it.</p>
<p>R. Researcher's questions (throughout the session)</p> <p>Clarity:</p> <ol style="list-style-type: none"> 11. Are there any categories, items, words or sentences that seemed confusing or ambiguous? Please underline vague items and rewrite them using other alternatives. 12. Is there any terminology that is puzzling to you? 13. Is there any item that is out of place? Please refer to the items that you think should be moved and locate their new place in the checklist. 14. Cross out the items that you think are repeated or unnecessary. 15. Are the checklist directions of use clear? Can you understand what is required from you? <p>S. Checklist completeness:</p>	<p>Teacher's answers</p> <p>Clarity:</p> <ol style="list-style-type: none"> 1. Categories are clear 2. No 3. P7 item No.b (teachers' needs section) is confusing (needs more explanation) 7. P.9 No. i : pedagogical approaches keep changing, so this item is unnecessary . No need for it in teachers guide 8. Yes <p>Completeness:</p> <ol style="list-style-type: none"> 2. P.3 No. h: Be more specific of what you are looking for. Reality of language use varies from one person to another. (needs more explanation) <p>P.4 N. b (More details)</p>	<p>Notes</p>

<p>5. Should more explanations be added? Which explanations can be added to improve the checklist? Where are the categories and items that need more explanation? (Suggest any missing items that you think are supposed to be included in the checklist.)</p> <p>6. Do you feel that you need more help in understanding the checklist?</p> <p>T. Difficulty /ease of use</p> <p>11. Can you use the checklist within a reasonable amount of time?</p> <p>12. Do you think that you could use this checklist without help?</p> <p>13. Do you feel that its use is challenging for ordinary teachers?</p> <p>14. Where there any difficult parts?</p> <p>15. Do you feel bored? Why?</p> <p>U. Language obvious errors:</p> <p>7. Please refer to any obvious errors such as misspelling, incomplete sentences, poor grammar or any other mistakes that you could notice.</p>	<p>P.6 No. b: Include the goals of learning English</p> <p>2. Identify what you are looking for: listing what you are looking for in all sections.</p> <p>Easy or difficult:</p> <p>6. Thoughtfully within 1-3 hours</p> <p>7. Yes</p> <p>8. No</p> <p>9. No</p> <p>10. No, but it gives you the sense that “I have a job to do”.</p> <p>Language problems:</p> <p>1. P.5 No. 3 a (author’s claims)</p> <p>2. P.8 No. c</p>	
<p>V. Behaviors to observe and record (Observations on specific items/parts of the checklist)</p> <p>7. Puzzled looks</p> <p>8. Long pauses</p> <p>9. Misunderstanding of some items, words or sentences</p>	<p>4. The puzzled looks and pauses were because the teacher was “comparing” the checklist items with what she sees in the textbooks she is using.</p> <p>5. No very long pauses</p> <p>6. No major misunderstanding probably due to the teacher’s post graduate degree.</p>	<p>Notes</p>

W. Unanticipated outcomes and comments	1.None 2. None 3.None	
X. Debriefing questions (at the end)	Answers	Notes
11. What are your general comments and recommendations of the checklist that you think will make it more practical and effective?	6. The checklist is good for the “evaluation of teaching materials”	
12. Do you think that other teachers will be interested in this checklist? Why/why not?	7. Yes, if “given a chance”, but some won’t be interested	
13. Would you use this checklist to select or evaluate your textbooks?	8. Yes	
14. What would you change in this checklist to make it better?	9. Only the previous comments	
15. What did you learn from using this checklist?	10. “It did refresh things” from my experience in teaching. I made her think “why you like something in a text book or why you didn’t like it”.	

Appendix H5: One -to-one Feedback (Teacher 3)

Teacher 3: Male, Master's degree in education, 3 years' experience, Omani

Thursday, 10/ 11/ 2016 from 10: 10 till 11:35

Y. Subject's (Teacher's) comments	Page and category or item No.	Notes
9. Comprehensive 10. I "need more involvement with the teaching materials"	No comments at this stage from the teacher apart from the general ones.	Reading the checklist took longer time for the teacher in order familiarize himself with the checklist.
Z. Researcher's questions (throughout the session) Clarity: 16. Are there any categories, items, words or sentences that seemed confusing or ambiguous? Please underline vague items and rewrite them using other alternatives. 17. Is there any terminology that is puzzling to you? 18. Is there any item that is out of place? Please refer to the items that you think should be moved and locate their new place in the checklist. 19. Cross out the items that you think are repeated or unnecessary. 20. Are the checklist directions of use clear? Can you understand what is required from you?	Teacher's answers Clarity: 1. Main category A p. 2 (Research-based aspects) is suggested to be replaced by curriculum theories or learning theories. 2. Precision in phrases and terminology (make them more clear. Use them as adjectives), for example pedagogical considerations 3. Nothing 4. P.3 item g: more details and examples on "aesthetic and emotional involvement" P.4 item i: Replace lives with "life styles" P.6 item b under students' needs: the items that guide the evaluator "some are specific and some are general". Some items are not clear "understanding key strategies" and repeating and recycling" which skills or inputs to repeat? 5. To high extent (suggesting guide, manual for use and for explaining some items) Completeness: 1. Use "terms from the literature" for "educators" ...to be understood by them - P.4, item c : other skills activation like Listening not only reading texts.	Notes

<p>A. Checklist completeness:</p> <p>7. Should more explanations be added? Which explanations can be added to improve the checklist? Where are the categories and items that need more explanation? (Suggest any missing items that you think are supposed to be included in the checklist.)</p> <p>8. Do you feel that you need more help in understanding the checklist?</p> <p>B. Difficulty /ease of use</p> <p>16. Can you use the checklist within a reasonable amount of time?</p> <p>17. Do you think that you could use this checklist without help?</p> <p>18. Do you feel that its use is challenging for ordinary teachers?</p> <p>19. Where there any difficult parts?</p> <p>20. Do you feel bored? Why?</p> <p>C. Language obvious errors:</p> <p>8. Please refer to any obvious errors such as misspelling, incomplete sentences, poor grammar or any other mistakes that you could notice.</p>	<p>9. Guide/ detailed description of the purpose of categories/ items used. Explanation of the process of categories and items selection-basis of selection-how did you reach this?</p> <p>Do the categories cover all the required areas!</p> <p>Easy or difficult:</p> <p>11. It will take time if good evaluation has to be done/ suggested that starting with “training sessions will help”.</p> <p>12. “with short training session+ guide) (no!!)</p> <p>13. To some extent, there may be specific items that need focus/ attention</p> <p>14. No</p> <p>15. No</p> <p>Language problems:</p> <p>2. P.8 No. item c Are instead of is!!!! Second opinion P.6 item b: not clear</p>	
<p>D. Behaviors to observe and record (Observations on specific items/parts of the checklist)</p> <p>10. Puzzled looks</p> <p>11. Long pauses</p> <p>12. Misunderstanding of some items, words or sentences</p>	<p>7. Little confused because of the phrases and terminology</p> <p>8. Pauses attempting to answer that specific item and isolate it from others.</p> <p>9. Only as discussed above.</p>	<p>Notes</p>

E. Unanticipated outcomes and comments	None apart from the need for longer time to read the checklist.	
F. Debriefing questions (at the end)	Answers	Notes
<p>16. What are your general comments and recommendations of the checklist that you think will make it more practical and effective?</p> <p>17. Do you think that other teachers will be interested in this checklist? Why/why not?</p> <p>18. Would you use this checklist to select or evaluate your textbooks?</p> <p>19. What would you change in this checklist to make it better?</p> <p>20. What did you learn from using this checklist?</p>	<p>11. Use terms from literature, start with workshop</p> <p>12. Some would be interested because it may help the teachers to be more critical about textbooks.</p> <p>13. Yes</p> <p>14. Comprehensive</p> <p>15. "That I need to know more about curriculum design"</p>	

Appendix J1: Small group presentation Slides

Teaching materials evaluation in English Language programmes: A Viable checklist



From <http://www.ing.ac.uk/learning/teaching-materials/evaluation/2010/>

1

The Role of Textbooks

- ▶ Course books are essential for students and teachers in English language programmes
- ▶ Most researchers agree that "it is widely accepted that the course book lies at the heart of any English language teaching situation (Sheldon, 1988; Hutchinson and Torres, 1994, cited in Iqbal, 2011)

2

Materials selection and evaluation

- ▶ Currently, the availability of English textbooks is not a problem anymore as there are hundreds of titles available in the market.
- ▶ However, users find difficulty in choosing the appropriate textbook.
- ▶ Colleges of Applied Sciences as an example

3

Materials definitions

Tomlinson (2012) defines materials development as "all the processes made use of by practitioners who produce and/or use materials for language learning, including materials evaluation, their adaptation, design, production, exploitation and research".

"Materials is a term used... to encompass both texts and language-learning tasks: text presented to the learner in paper-based, audio, or visual form, and/or exercises and activities built around such texts" (Harwood, 2010)

4

Materials Definitions

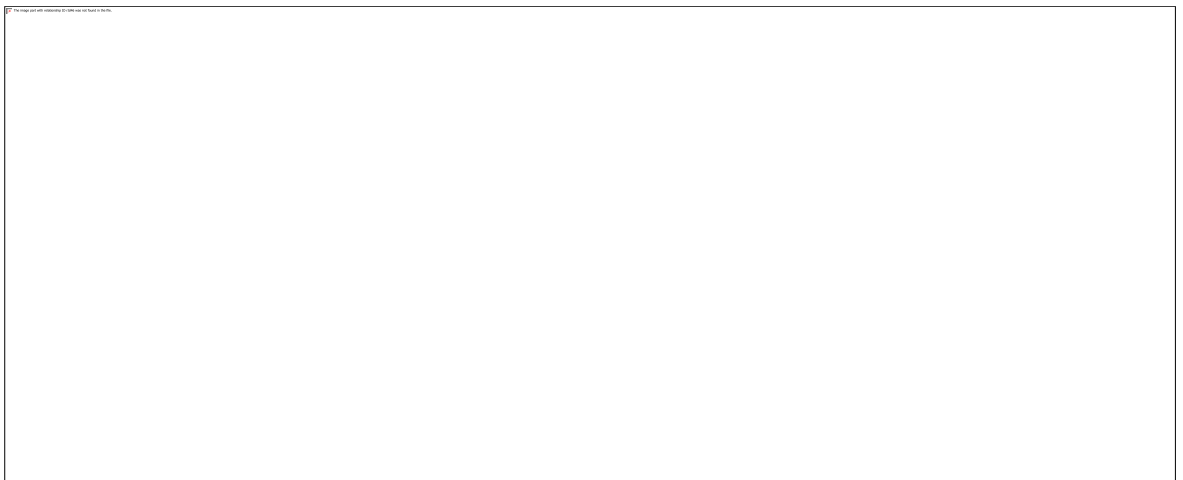
- ▶ Tomlinson (2001, 2013) has his own definition to materials that "include anything which can be used to facilitate the learning of a language.
- ▶ McGrath (2013): "A textbook, produced by a commercial publisher/ Commercial materials that are not provided as part of the textbook package/ Teacher-prepared materials/ Any other visual or auditory means used in the classroom by teachers or learners (facial expressions, gestures..."

5

Materials development models

- ▶ Ralph W. Tyler curriculum development Model (1949)
Goals and objectives/ content/ learning experiences/ evaluation
- ▶ Taba curriculum development Model
Statement of goals/ needs analysis/ Formulation of objectives/ content selection/ content organization/ Learning experiences selection and organization/ Evaluation (to check if the material achieve the intended learning outcomes)

6



Appendix J2: Small group review questionnaire (for teachers and experts)

Questions (to identify trouble spots)	Participants feedback
A. Effectiveness	
1. Does this checklist help you to evaluate and understand the textbooks you are using? How?	
2. What are the weaknesses and the problems that you noticed in the checklist?	
3. How can these problems be solved?	
B. Usability and Practicality	
1. Do you find the checklist easy to use? In what way?	
2. Do think you will need previous knowledge or instructions to use it?	
3. What are the instructions or knowledge that you think will facilitate the use of the checklist?	
C. Appeal	
1. Do you find this checklist interesting and attractive to use? In what way?	
2. What did you like most about the checklist?	
3. What did you dislike about it?	
4. Please write any suggestions to make this checklist more ingesting for its potential users (materials evaluators: teachers, coordinators, experts).	

Appendix J3: Small group review feedback (teachers)

Questions	Teacher 1	Teacher 2	Teacher 3	Teacher 4	Teacher 5	Teacher 6
A. Effectiveness						
1. Does this checklist help you to evaluate and understand the textbooks you are using? How?	Yes, it does. The questions set make me aware of things which I have not about previously.	The checklist covers many relevant questions in general terms. Visually the checklist page could and should be more easy to recognize at a glance.	Yes, it helps me understand the needs of an ESL textbook and have a standard by which to judge them	Yes, it helps analyze the components needed to make course material become functional	Yes, it does. It highlights critical issues related to T.B. evaluation	Yes, as a guideline to make changes and recommendations for future curriculum changes
2. What are the weaknesses and the problems that you noticed in the checklist?	Sometimes, too many things put under one item.	Sub-headings on checklist could/should be more block by block clear from first glance.	I would change available/not sure/not available to “on a scale from 1-3 how much do you agree”	Some of the participants the scoring confusing	The questions on relevance to the course (within CAS) could be broadened	It must be general not based a particular group. As a tool to evaluate materials is fine, but as a tool to evaluate a group of students with a particular book is flawed.
3. How can these problems be solved?	By designing more elaborate one (checklist)	Sections: contents of books/ classroom interactions/ Secondary resource learning <i>Clear sub-headings (for the user from first glance)</i>	Same as above	More consultation with educators	The questions on the relevance suitability of the coursebook viz. the foundation course and the assessment tasks	Use the checklist to evaluate materials without a reference. Just as teaching materials only!
B. Usability and Practicality						
1. Do you find the checklist easy to use? In what way?	Yes, I do	It is clear Colours are clear (sections)	Yes, its very simple and intuitive	Yes. The three choices helped. The “don’t know”	Yes, it is simple with only 3 options	Yes, it is comprehensive.

		Print is too small to answer adequately		section stopped any confusion		
2. Do think you will need previous knowledge or instructions to use it?	No, I do not	yes	Just experience as a teacher	The longer teaching experience helped understand the questionnaire	Not really, The instructions are clear.	No
3. What are the instructions or knowledge that you think will facilitate the use of the checklist?	To use simple and straight forward ideas	A verbal and visual explanation of the checklist parts and terms.	Brief instructions on how to fill out the survey and its purpose	Those that relate to the practicality of materials provided	Knowledge: → experience only the coursebook being evaluated, formal TEFL training	To be familiar with the materials you are about to evaluate
C. Appeal						
1. Do you find this checklist interesting and attractive to use? In what way?	Yes, I do. It is interesting and stimulating	Excellent and relevant concepts included shows real knowledge of teaching/ classroom process	Yes, it is easy and well layed-out	Interesting yes. Attractive? More pictures maybe	Yes, It has a clear design and the prompts are specific.	It looks clear and user friendly
2. What did you like most about the checklist?	It is extensive and clear	It looks so neat and uncluttered	It is well organized, clear instructions	Its linear layout	Close reference to key issues regarding textbook evaluation	The question of teacher/ student relevance in dealing with specific teaching materials.
3. What did you dislike about it?	Sometimes it is not clear	1.Small print 2. Lengthy and wayward explanation descriptions	Small font	No pictures	None	The lack of not addressing the impact that language carrier all the cultural aspects of L2
4. Please write any suggestions to make this checklist	To use as less items as possible	1. Large print 2. Secondary sheet for	Change to a scale system (i.e. 1-3 or 1-5, etc...)	Maybe yes- no- don't know instead of available- not available	None other than mentioned in 3 on page 1.	Language cannot be separated from culture. I think it is important to address the

<p>more interesting for its potential users (materials evaluators: teachers, coordinators, experts).</p>		<p>longer responses</p> <p>3. Clear, functional and useful sub-headings</p> <p>4. Much shorter topic sentences for each question/ area</p> <p>5. Overall organization must be well thought out not just considering 3 components of checklist</p> <p>Goal <u>assessment</u> areas (<i>sub-headings should apply to the users (teachers)</i>)</p> <p>6. Headings are good but do not stand out enough.</p>				<p>language-culture dichotomy</p>
---	--	---	--	--	--	-----------------------------------

Appendix J4: Small group review feedback (Experts)

Questions	Participants feedback	
A. Effectiveness	Expert 1	Expert 2
1. Does this checklist help you to evaluate and understand the textbooks you are using? How?	In a general sense, the checklist does cover several aspects of concern that can theoretically be analyzed and reported on by expert members.	Yes, it helps evaluate materials as it covers most important aspects of evaluation.
2. What are the weaknesses and the problems that you noticed in the checklist?	The first impression that was received from the title page was the subtitle “Quick Evaluation Checklist”. This title is misleading on several aspects. First of all, the total checklist is 10 pages in length, so it hardly constitutes a quick evaluation. This, in turn, could lead to suspicion, mistrust, and katzenjammer amongst the expert individuals or respondents who might otherwise be unhappy to make use of the checklist. Also, the reliability of responses may be made suspect, as many potential participants may give up or give inaccurate responses due to the misrepresented time anticipated to finish. Another weakness is the numerical scale assigned to measure differences in responses that, I feel, could be more accurately represented by verbal descriptions on a wider range of values or response points. Finally, there are several grammatical and linguistic errors that are made which will also contribute to the reluctance of participants to finish, and therefore also potentially affect the broad area of validity of the checklist and reliability of responses.	A few errors in word order
3. How can these problems be solved?	First of all, a rewording of the title is crucial to more adequately represent the task(s) that you are most likely expecting your participants to accomplish. As a suggestion, a detailed evaluation checklist is more informative in conveying the purpose, as well as the purview, of your study. Secondly, the scale values to be used for this checklist, given the probable community of experts you wish to tap, could likely be more effective if converted or expanded to include	Proof-reading

	<p>more variations and degrees of possible responses.</p> <p>Finally, there should be a very scrupulous review given to the grammatical and linguistic errors that currently exist in the checklist items themselves.</p>	
B. Usability and Practicality		
1. Do you find the checklist easy to use? In what way?	<p>It's easy to respond to simply because there are only three values from which to select, however, looking at the entire document, it is actually more complex to squeeze the many opinions and attitudes held about the questions asked within the three rather ambiguous scale values. As noted earlier, the overall length of the checklist, combined with the somewhat opaque rating scale, discourages or hinders the probability of a completed checklist, and therefore of an acceptable response rate.</p>	<p>Yes, it is categorized into clear dimensions</p>
2. Do you think you will need previous knowledge or instructions to use it?	<p>As a case in point, the very phrasing of this question “<i>Do think you will need previous knowledge or instructions to use it?</i>” may hinder the willingness of potential respondents to get involved, as there exists a grammatical error in the question. Answering the question as intended, the details involved in the scenarios suggested very strongly suggest that the ideal respondents should be experts, and they should have the pre-requisite amount of knowledge to be able to answer in as comprehensive a manner as the checklist suggests to be done. Unfortunately, based upon these same backgrounds, they may conclude that you do not have enough details or credibility to legitimately pose questions about these issues, as reflected by the superficial flavor in the construction of the checklist items and questions.</p>	<p>Yes, to understand it previous knowledge is necessary</p>
3. What are the instructions or knowledge that you think will facilitate the use of the checklist?	<p>Possible examples of suitable texts or books should be included with the transmission of the checklist.</p>	<p>You have to give instructions to the evaluators on how to use it</p>
C. Appeal		
1. Do you find this checklist interesting and	<p>The comprehensiveness that the checklist suggests by its mere length is compromised by the relatively</p>	<p>Yes, it is practical</p>

attractive to use? In what way?	incomplete and inaccurately posed questions and scale values offered within the checklist itself. Regrettably, this checklist does not match the complexity that its ideal respondents would most likely find attractive in that it does not accurately reflect all of the different opinions they may hold about the subject matter.	
2. What did you like most about the checklist?	The detail of the criteria in each section or component, as well as the selection and total number of possible components or evaluation areas that are listed on each page.	It covers the most important issues related to materials evaluation.
3. What did you dislike about it?	I am obliged to say that, taken as a realistic feedback device, this checklist needs to be significantly improved. The word <i>dislike</i> is not appropriately used in this context. Whether or not a respondent likes or dislikes it is not as important as the rating of all of the qualities that it intends to measure. So, there is no appropriate venue to express personal feelings at this level.	It is a little bit long
4. Please write any suggestions to make this checklist more interesting for its potential users (materials evaluators: teachers, coordinators, experts).	Try to envision the kinds of respondents who you would ideally like to have for your checklist, and try to dovetail the anticipated level of expertise that you desire with language and internal conceptual framing that that community would most logically respond to. Further, I would recommend looking at our suggestions in the previous sections provided for feedback, to improve the internal or face validity of your checklist.	It is a little bit long so you can shorten it

Appendix J5: Small group observational log (teachers)

Teachers	Completion time	While use problems	comments by participants
Teacher 1	54 minutes	Is the checklist suitable for all levels	
Teacher 2	33 minutes	Scoring criteria is not clear The size of rows and columns	Simplify the terms used
Teacher 3	22 minutes	The comments were written on checklist	
Teacher 4	40 minutes	Same as teacher 3	
Teacher 5	40 minutes	Is the checklist for a whole package or just the main textbook (coursebook?)	Reading/writing texts do not match the final exams/ assessment
Teacher 6	25 minutes		The textbooks are high above students' level/ are beyond their understanding

Appendix J6: Small group observational log (experts)

Experts	Completion time	While use problems (effectiveness, practicality and use)	comments by participants
Expert 1	Start: 9: 34 End : 10:04 Comp. Time: 30 minutes	The term literature review needs to be thought of	1. It is a long checklist 2. It covers most aspects
Expert 2	Start: 9: 34 End: 10: 10 Comp. Time: 40 minutes	“rationale of the construction of the checklist items”	1. “it is very lengthy” 2. Suggested focus group/ interviews to discuss the items in details.

Appendix K1: Instructions for field testing sessions

A. Rationale for the Study

First, a conceptual framework is developed by the researcher and validated by four experts to be used as a starting point for the checklist design and its main categories. Based on the framework (attached below), two main checkpoints were established: literature or research and setting needs. The main categories in research were based on second language acquisition principles (what to teach), teaching principles (how to teach) and ELT Curriculum design (the way what and how are organized). The main categories have several items in “what to look for” column to evaluate the teaching materials. To reach this stage, the checklist is revised through 5 cycles of formative evaluation: developer screening, expert review, one-to-one review, small group review and finally field testing.

Field testing includes using the checklist to evaluate the following materials:

- The materials that you already taught or the materials that you are teaching this semester. Please make sure that you have the whole package including the workbook, the CDs and the teachers manual.

B. How can users inform the researcher about the problems they encountered while using the checklist?

- At this stage of checklist testing (field test), the purpose is to know how users will be using the checklist on their own without any help apart from the instructions provided in the emailed documents.
- While the session, the participant writes any questions or notes in the sheet provided to tell researcher about these problems
- The participant will also answer the feedback questionnaire at the end of the session
- All your comments, notes, questions and feedback will be considered when revising the checklist in this developmental stage. So, please include anything that you think will help to make this checklist a user-friendly tool for evaluating teaching materials in the English Foundation Programmes.

C. How to use the checklist

The checklist is divided into two main parts. Part A: the quick evaluation and part B: the close or the detailed evaluation. Each one is in separate sheet for ease of use.

1. Go through the checklist to become familiar with its content

2. Bring the materials that will be evaluated (materials you taught last semester or you are teaching this semester)
3. Evaluate the materials against the items in the quick checklist first.
4. If the evaluation score is above 80% you will automatically go to the detailed evaluation sheet
5. Repeat the evaluation for all the components in the detailed evaluation sheet.
6. If the total score is 60% or above you can select or reuse the materials for your students and if less reject and look for other appropriate materials

Please note that:

- You can quit the session at any time if you feel uncomfortable for any reason
- If you have any questions, please ask them before the session or after the session as the researcher/ observer will not be able to answer them during the process of field testing.
- Please make sure to record the start and the end time of the session as it is very important for the researcher to know the time needed by users to finish the checklist.
- The conceptual framework of the checklist basis and sources is included with this summary to help you understand the checklist and its main headings and sub-headings.

Many thanks for your participation and cooperation

Appendix K2: Field testing questionnaire

1. General information:

- a. Name (optional):
- b. Gender:
- c. Education:
- d. Teaching Experience:
- e. Nationality:
- f. Date:
- g. Start time:
- End time:

2. Feedback questions on implementation (at the end of the session)

Question	Feedback
<ul style="list-style-type: none"> • Did you have problems approaching or start using the checklist? 	
<ul style="list-style-type: none"> • Did you have problems while using the checklist? 	
<ul style="list-style-type: none"> • What kind of problems did you face? 	
<ul style="list-style-type: none"> • Is additional training needed on how to use the checklist? 	
<p>Are more guidelines needed for the use of the checklist?</p>	
<ul style="list-style-type: none"> • Does the checklist satisfy the users' need (in your institution) for evaluating teaching materials? How? 	

Appendix K3: Participants Field testing feedback

Questions	Participant 1	Participant 2	Participant 3	Participant 4	Participant 5	Participant 6	Participant 7	Participant 8	Participant 9	Participant 10
Did you have problems approaching or start using the checklist?	No	I couldn't read the 2 nd checklist when the file opened (had to zoom)	No	There was no problem.	No	Weak internet signal Old Excel version Understanding few bullet points in the checklist . (but they were clarified by the researcher.)	No	Not so much, but I did need some help understanding some of the questions, mainly to give a more accurate answer.	At the beginning, I think it was clear and there was no issue with the sheet.	No problems
Did you have problems while using the checklist?	Yes. Changing between pages	Yes	Yes	Some problems occurred.	When I was I going through the detailed checklist, my answers were much affected by my own experience of teaching the series of the book	No	Yes,	Maybe it was going to be slightly better if I had more options to choose from rather than the 3 options provided (Yes, No, Not sure). However, the research	Yes, I felt like I needed more options to add the accurate. I don't really have the exact answer for questions like the price of these books.	None

					used in the session. Two minds, what I have gone through and what I want to see in a textbook if I had to select one			er has surely chosen these for specific purposes related to the research design.		
What kind of problems did you face?		Some questions were unclear e.g. “follow the description from listening texts”. Some Qs impossible to answer as they contained multiple Qs * “accents <u>and</u> real conversations” “ Tips for speaking <u>and</u> writing”	Some columns consistency	A major problem was on (Teacher’s Needs) and its statements/phrases. I think that most of them were not related to teacher’s needs (preferences, beliefs, personality / identity etc.). Also combining two different terms (teaching and learning) in one statement was confusing.	Some typos need to be checked and corrected through out. Consistency should be maintained of some terms (e.g., L2 for English language). Participants should be told at the beginning of the session	---	1. Some statements were not shown as complete statements on the excel sheet.	I would’ve preferred a likert-scale continuum to answer some of the questions	There is no scale to add the exact opinion The content, I felt that there are so many questions and personally speaking I don’t really think that teachers will spend all that amount of time in answering them unless a specific session is conducted and prepared only for	None

					about the type of evaluation (i.e. pre- and/or post).				this purpose	
Is additional training needed on how to use the checklist?	No	No	No. They are very easy	If the checklist items are modified with more specific items, there will be no need for extra training.	No	No	No	No, I think most of it was manageable.	Yes, I think there is a need for that because not all teachers are really aware of curriculum design and Also, teachers need to be informed about the importance of this checklist in order to answer it.	No, but teachers should know the materials beforehand
Are more guidelines needed for the use of the checklist?	A brief outline to the contents/ topics asked	No	No.	I think the researcher needs to clearly define some pedagogical terms such as course, syllabus, and curriculum	No	No	No	Yes, in some cases, but only at the start.	Yes, I think there should be a clear guidance and instructions informing the teachers to choose the target book	Not really
Does the checklist satisfy the users' need (in your institution) for	Yes, helpful.	We need to evaluate if material	Yes. Clear options.	After more modification (especially	Yes, to a very good extent	I think yes. The checklist	Yes I think it include	I think YES. Using this	I think yes, However, I suggest	Yes, it does. It evaluate

<p>evaluating teaching materials? How?</p>		<p>facilitates the move from general to academic English</p>	<p>Simpl e state ments . Under standa ble langu age and instru ctions</p>	<p>the section of Teacher’s Needs), I believe the checklist could be used for evaluating teaching materials. However, the checklist needs to organize items according to their categories (layout and design, curriculum, language skills, language content, topic content, activities, methodology, learner’s needs, teacher’s needs, institutional needs)</p>	<p>The current study only focused on the Course book content level of analysi s but future researc h could move further and include other levels of analysi s (e.g., consum ption and product ion).</p>	<p>provides valid points to consider when choosing a text book for the Foundati on program me or any other ELT program me.</p>	<p>s the main points to be considere d when evaluati ng any material s.</p>	<p>checklist can help us narrow down our choices when choosin g the best teaching material s and textbooks. However, I think it’s also a bit long and some question s might be more related to teaching material s (supple mentary, rather than textbook s).</p>	<p>adding couple of questions targeting the ability of teachers to use the technology requireme nts- in the text book- to activate the lesson</p>	<p>es the differen t element s such as, visuals, teacher ’s notes, supple mentar y materia ls etc.</p>
--	--	--	---	---	---	--	---	--	---	--

Appendix K4: Field testing observational log notes

Observation/ thought	Participant 1 (Coordinator)	Participant 2 (Teacher)	Participant 3 (Teacher)	Participant 4 (Coordinator)	Participant 5 (Expert)	Participant 6 (Teacher)	Participant 7 (Teacher)	Participant 8 (Teacher)	Participant 9 (Coordinator)	Participant 10 (Teacher)
Time	12: 38 01: 15 37 minutes	12: 38 01: 19 41 minutes	10: 34 11: 50 1 hour 24 min.	10.35 11.50 1 hour 23 min.	11:01 11:42 41 minutes	10:15 10: 51 36 minutes	1:34 1: 57 24 minutes	4:48 5: 36 48 minutes	10:17 10:41 24 minutes	End of questions 11:03
Performance			<p>Quick checklist: same as teacher 4</p> <p>Detailed checklist The size of the second page (sheet) should be in appropriate size</p> <p>Authentic materials: consider the use of this term : a section on what is authentic</p> <p>Second language acquisition principles Items b. And: in</p>	<p>Quick checklist Few spelling and layout issues (etc.), Spelling of a word (detailed) Missing word due to the change in the checklist (6. Item 1) Price? Does not know</p> <p>Detailed checklist Sheet 2: year of publication: capitalization</p> <p>Teacher's needs:</p>	<p>-The font size of some options from the drop-down list is very small</p> <p>-The expert's excel sheet was not compatible with the version used to design the checklist</p> <p>-The participant's previous experie</p>	<p>Simple typing mistakes may make it difficult for some users (e.g. detailed</p>	<p>- Checklist font size when the detailed checklist first opened. The participant was directed to the place where he can change the view of the checklist.</p> <p>-Some missing items due to the small height of some rows (e.g.</p>	<p>The quick checklist The price item in the quick checklist raised some concerns for the teachers as they do not know how much the material's cost. This participant searched for the price and inferred that it is about 120 OR per</p>	The performance was fine and smooth	

		<p>the options the font size is very small</p> <p>Curriculum Design</p> <p>All the columns need to be fixed...problem with the inconsistency of options (needs to be fixed)</p> <p>Curriculum Design:</p> <p>No b. items 2 (in the teachers' book instead of textbook)</p> <p>Needs Analysis</p> <p>Students' needs: b (1) a No. 3, No. 4 (themes) spelling (Students) spelling</p> <p>To "communicating</p>	<p>Item 3: compatible with assessment (under the course not teachers' needs)</p> <p>Also item 5... In relation to the course implementation (yes) but other than that it is not related to the teachers' needs...e.g. current and up to date</p> <p>Other items like "flexibility in choosing materials" are more suitable (consider teachers' preferences for.....<u>in the feedback</u> answers</p>	<p>nce with the teaching of the same materials presented in the session affected their evaluation and the selection of items in the checklist.</p> <p>-More details were needed on the teachers' and the learners' role: in what way</p> <p>-Row 43 question "to what extent instructions are explained to teachers? In textboo</p>	<p>SLA Principles "a" items 1 and 2). The participants could solve these problems by himself when the cell was activated. Very easy and smooth use apart from a technical problem that is due to coding in the answers where there were some answers already in the checklist, so the participant had to select the appropriate. The problem</p>	<p>student and he thought that it is expensive. So, involving the teachers will make them appreciate the value of the materials pedagogically and financially.</p> <p>CDs not useful (unreal conversation)</p> <p>Exams suitable for students level, but not students' needs</p> <p>Needs' of the students are different now</p>	
--	--	---	---	---	---	--	--

			with other cultures”		k or teacher’s book? Needs clarification.		was explained to the participant and told that I was not intentional but a problem with the excel sheet.		
			<p>Teachers’ needs</p> <p>Item “d” and A No. 5 (do materials)</p> <p>Institutional needs</p> <p>Do materials</p> <p>Item “b” to what extent can materials</p>	<p>(research skills)What if only two units have research-based activities ? Yes/ no or not sure?</p> <p>Font is not consistent: 12 would be good for all</p> <p>Ministry of Higher Education: capital letters</p> <p>Satisfaction surveys for students yes</p> <p>For teachers No</p>	<p>Consistency in the use of terms e.g. second language and English language</p> <p>-Row 50: some typos: themes and incomplete sentence</p> <p>- Teachers’ needs B:item 2 and 4 (follow descriptions) it needs to be clarified</p>				

					<p>- Context is very important: what is the difference between learners' context and teachers' context?</p> <p>- Teachers' context: perceptions of teachers, cultural background and language teaching background (novice vs. experienced)</p> <p>So the question would be</p>				
--	--	--	--	--	--	--	--	--	--

					more appropriate if it is rephrased as “to what extent are materials suitable to the learning context / teaching context instead of teachers’ or learners’ context . -Level needs to be specified (of language proficiency)				
Behavior	The session went very quietly with some consultation of the text books	The session went very quietly and participant looked at the books	The participant discussed lots of issues and later wrote his comments in the evaluator	The participant focused on details and continued the evaluation after	- Making the detailed checklist hidden may confuse the evaluator	The need to discuss the checklists loudly to check the	<u>Calm</u> and enthusiastic	<u>Consulting the textbooks</u> for the availability of some items	<u>Quiet throughout the session</u>

		several times	s'' notes section	referring to some items in the checklist	ors thinking that there is only that short checklist which does not cover all the items in conduct ing a good materials' evaluation. So, making both checklists apparent would be more easier for users	meaning of some items ...suggesting that using the checklist among a group of teachers will be useful				
Notes and thoughts	The automatic move to the detailed evaluation checklist did not work. The cause can be the internet	-I have to provide a hard copy with one column and the equation calculation	Researcher's thoughts -Should second sheet be hidden? -Even small things like font size do matter -New changes/corrections may cause new mistakes ...so there is always something to improve... Also, the time allocated for corrections is very important... more time less mistakes. -Every teacher is a living experience with new thoughts, new	The drop-down list of the three options (yes/no / not sure) did not work in the expert's	- Appropriate size mean... -SLA principles... "P" item 2: share (does it mean to	The prices are not always known to the teachers ...but price has to be considered when selecting	Some participants discussed some of the items in the quick checklist suggesting their interest in its	He only needed a short explanation of the detailed content and main categories The item about price is		

<p>connect on, so this has to be fixed.</p>	<p>where the evaluators can do it manually. Provide a sheet for notes / problems while use - Teachers evaluated materials that they are familiar with. So, new materials may take longer time to evaluate. Also, a hard copy may take less time to complete</p>	<p>ideas and new suggestions.... -At certain point, participants had to be reminded that their focus is the checklist not the materials at this session. -The two participants who are conducting their higher education degrees have more comments on the content of the checklist more than the use due to their current interests in this area.</p>	<p>computer</p>	<p>talk about? - Teaching principles... “a” item 1: meaning of “throughout the materials” ... better to say in almost all units -The importance of the evaluator’s notes column : for the evaluators and for the other users and stakeholders to include any concerns arisen from the checklist use or the materials</p>	<p>g teaching material s... Researcher’s thoughts The excel sheet needed more preparation because some participants may not be familiar with excel, so they could not deal with the sudden problem s. This point is inferred because of the repetitive cycles DBR which allow the researcher to connect things through the process</p>	<p>content and items Quick checklist: all the evaluation scores were above 80%, but in the detailed checklist the scores are lower Researcher’s thoughts Some participants who pay attention at every detail took them longer to finish the evaluation. It is interesting to see how different participants</p>	<p>also raised... Though its importance in comparing the materials and in informing teachers about the financial value of the materials they teach which may provoke interest and criticism. Considered all the levels in the series with focus on level A. This can refer to the usability of the checklist with different levels Researcher’s thoughts Should the suppleme</p>	
--	---	--	-----------------	---	--	--	---	--

					<p>ls' evaluation (e.g. Notes as the use of item in only one unit of the textbook: yes or no!)</p> <p>-2. Teachers' needs: "a" item 1: current and up to date: what is it? ..Teaching materials</p> <p>"b" item 6: personalized activities: mean: talk about themselves!</p>	<p>of data collection. Also, as mentioned before the appearance of problems result in new solutions such as providing a word document checklist which in return led the researcher to compare the excel sheet version and the word document version</p>	<p>deal with the checklist and its use. The evaluation and selection of teaching materials need collaboration as well as a thorough instrument like the developed checklist. A committee that consists of the <i>details detector</i> and <i>quick decision makers</i> can solve lots of problems related to teaching materials in the Foundation Programmes.</p>	<p>ntary that are provided by the teachers in the college be part of the evaluation? No, because the focus is on the published materials. Though the in-house materials was evaluated by 2 experts in small group evaluation, which dominate that the checklist can be used in that regard.</p>	
--	--	--	--	--	--	---	---	---	--

Appendix L1: Teaching materials evaluation checklist prototype 1

Text book title:					
Author:					
Level:					
Year of publication:					
I. Impressionistic evaluation (first glance) checklist					
The criteria	Their sources	What to look for?	Available	Partially Available	Not available
1. Practical considerations	McGrath (2002, p.33)				
		Multi-level			
		All components are available			
		Affordable			
2. Support for teaching and learning					
		Additional components			
		Teachers' book			
		Tests			
		CDs			
		Suitable for self-study			
3. Context relevance					
		Suitable for the course			
		Length of course			
		Aims			
		Syllabus			
		Exams			
		Suitable for learners			
		Age			
		Level			
		Cultural background			
		Suitable for teachers			
4. Likely appeal to learners		Layout			
		Visuals			
		Topics			
		Suitable over the term (unlikely to date)			
The textbook (s) that passed the impressionistic evaluation can be evaluated in details using the following close evaluation checklist:					
II. Close evaluation checklist					
The criteria	Their sources	What to look for?	Available	Partially available	Not available
A. Research- based aspects (from literature reviews)					

1. Second language Acquisition principles					
a. Materials should help the learners to develop confidence (involving them in tasks which are challenging but achievable)	Tomlinson and Masuhara survey review (2014)	Activities ask students to think and which “do not provide answers all the time”			
b. Materials should require and facilitate learner self-investment (through...responsibility of making decisions...make discoveries about language themselves)	Tomlinson and Masuhara survey review (2014)	Activities that help the learners to make discoveries about “how the language is used” and that ask them to discuss their findings.			
c. Materials should expose learners to language in authentic use (a rich and varied input which includes unplanned, semi-planned and planned discourse)	Tomlinson and Masuhara survey review (2014)	Use of authentic texts like newspapers and magazines Use of longer texts with less editing as much as possible Use of different accents and real conversations			
d. Materials should provide the learners with opportunities to use the target language to achieve communication purposes... to develop strategic competence	Tomlinson and Masuhara survey review (2014) McGrath (2002, p.47)	-Content that focus on “meaning and form” -Activities that “encourage communication between students” such as debates, making conversations and instructions how to speak and write effectively Activities that allow “opportunities for students to express their own meaning in their own words”			
e. Materials should take into account that the learners differ in affective attitudes.	Tomlinson and Masuhara survey review (2014)	Texts that can “amuse, excite and stimulate” students			
f. Help the learners to develop cultural awareness	Tomlinson and Masuhara survey review (2014) Adaskou, Britten & Fahsi, 1990 (cited in McGrath, I. 2002, p. 211)	Texts that consider other “cultures and countries” -Aesthetic aspect of culture: media, cinema, music and literature -Sociological aspect of culture: nature of family, home life, interpersonal relations...work, leisure, customs and institutions -Semantic aspect of culture: food, clothes and institutions -The pragmatic aspect (sociolinguistic): background knowledge, social skills...			
g. Materials should maximize learning potential by encouraging intellectual, aesthetic, and emotional	Tomlinson,B	Texts and activities that challenge learners to think analytically and creatively			

involvement which stimulates both right and left brain.	(2011 cited in Tomlinson. 2013, Pgs. 28-30) Tomlinson and Masuhara survey review (2014)				
h. Reflect the reality of language use	Tomlinson, B(2013) Tomlinson and Masuhara survey review (2014)	Use of language as a Lingua Franca (foreign accents and non- native speakers conversations) Activities that encourage the “use of language outside classroom)			
i. Materials should cater for the needs of all learners	Tomlinson and Masuhara survey review (2014)	Representing all types of lives: rural, urban and simple lives not only “middle-class, travelers and well-educated”			
j. Materials should help learners after course to develop “autonomous learning”	Tomlinson and Masuhara survey review (2014)	Activities that teach “real life strategies, skills” e.g. portfolios			
k. Materials should help learners to personalize their learning	Tomlinson and Masuhara survey review (2014)	Provide activities for learners involvement through asking for their’ “views and opinions”			
2. Teaching Principles and pedagogical factors					
a. Learners should have increasingly spaced, repeated opportunities to retrieve and give attention to wanted items in a variety of contexts.	From Nation/Macalister (2010)	Studies that support the effects of repetition on learning (Kachroo, 1962; Saragi <i>et al.</i> , 1978) and the levels of processing theory (Craik and Tulving, 1975). Check the course books to make sure that they provide repetition of the important items of the language several times			
b. The teaching of language items should take account of the most favourable sequencing of these items and should take account of when the learners are most ready to learn them.	Nation/Macalister (2010)	We still have only an elementary list of sequenced grammatical items to guide teaching and do not have easily applied tests to indicate the learners’ stage in the sequence of development (Pienemann <i>et al.</i> , 1988).			
c. The course should help learners make the most effective use of previous knowledge.	Nation/Macalister (2010)	A few questions at the beginning of the texts are designed to stimulate relevant knowledge			
d. The items in a language course should be sequenced so that	Nation/Macalister (2010)	Teaching all the parts of the body together, teaching a range of			

	items which are learned together have a positive effect on each other for learning and so that interference effects are avoided.		colours together, and teaching the series of numbers together...Some words are more frequent than others, so should be taught first			
e.	Learners should receive helpful feedback which will allow them to improve the quality of their language use	Nation & Macalister (2010)	-Regular opportunities for careful language production -Realistic list of aspects of language use that learners can be encouraged to monitor -Information gap or opinion gap activities which encourage peer negotiation -Regular use of an informative and acceptable marking system for written work			
f.	Successful instructed language learning requires extensive L2 input	Ellis, R. (2005)	Maximize use of L2 in the classroom Opportunities to receive input outside the classroom Extensive reading through graded readers And training on how to use available resources			
g.	Successful instructed language learning requires opportunities for output	Ellis, R. (2005)	Increasing oral and written tasks			
h.	The opportunity to interact in the L2 is central to developing L2	Ellis, R. (2005)	Small group work tasks (provided that learners use L2) Role play activities			
3. The ELT Curriculum and methods						
a.	The method or pedagogical approach of the materials is made clear to the users	Richards, J.C. & Rodgers, S. (2014)	-The author claims on the cover of the textbook, the introduction, or the teachers' manual are noticed within the textbooks. -The objectives are explained <ul style="list-style-type: none"> • -The syllabus • -Learners role • -Teachers role • -Materials role 			
b.	Procedures					
	o The use of Pedagogical activities is well explained	Richards, J.C. & Rodgers, S. (2014)	Detailed teachers' manual on how to use every activity and its purpose.			

<ul style="list-style-type: none"> ○ Procedures and techniques in giving the feedback on the activities to the learners are explained 	Richards, J.C. & Rodgers, S. (2014)	Instructions for the teachers on how to inform learners on various activities			
B. Setting-based Factors (needs analysis)					
1. Students' needs	Qualitative data (brainwriting sessions) from 3 different proficiency levels				
a. Materials should help learners accomplish their goals from learning English language	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Through inclusion of topics on self-development -Through inclusion of texts and conversations that promote communicating with other cultures and mutual understanding 			
b. Materials should include all language skills and items that are desired and specified by learners	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -More focus on the 4 skills : Speaking, listening, reading and writing as well as grammar, pronunciation, and vocabulary -Providing sufficient activities on dictionary use, dictation, and spelling games -Identifying and focusing on difficult words and topics -Learning how to form questions -Understanding key strategies in the textbook -Repeating and recycling -Learning words through writing 			
c. Materials should consider that learners differences in their study habits, learning strategies and styles	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Incorporating the appropriate and different strategies and techniques in delivering the language -Utilizing illustrative drawings, pictures, maps and infographics -Utilizing the available sources like internet, mobile phones, newspapers and magazines -Providing entertaining and educational tools like songs, short films and documentaries 			

		<ul style="list-style-type: none"> -allocating enough time for students to speak in class and be involved in teaching and presenting parts of the lessons to the class -Providing more students' teachers' interactions 			
2. Teachers' needs	Qualitative data (brainwriting sessions)				
a. General characteristics of a good teaching materials (textbooks)	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -The teaching methods of the materials are current and up-to-date -The assessment is compatible with the teaching materials -The materials are well organized -There topics are diverse -The content is interesting -The teaching materials are culturally appropriate -They are easy to adapt and edit -The materials accomplish the course objectives -They have tests to show students' progress -The materials require students' involvement -Materials incorporate current trends like gamification, critical thinking skills, research skills and e-learning tasks 			
b. The appropriate content and language items for listening and speaking	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Listening for specific information -Mini-presentations -More listening activities (3 minutes or longer) -Dialogues and turn-taking -Personalized speaking activities -Follow description from listening texts -Prompted and cued speaking -Note-taking skills 			
c. The he appropriate content and language items for the reading and writing skills	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Skimming and scanning skills -Looking for main idea in reading -Graded reading activities -Reading for fun activities -Reading comprehension activities -Paraphrasing tasks -Different genres of writing (essays, letters, emails...) 			

		<ul style="list-style-type: none"> -Guided writing activities -Free writing activities -low level sentence structure activities -Writing paragraphs and essays -Punctuation activities -Using connectors to join clauses 			
d. The appropriate content and language items for (vocabulary, Grammar and pronunciation)	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Vocabulary exercises -Soft copy of vocabulary list -Using idiomatic language -Using synonymous and antonyms -Verb tenses tasks -Grammar in context activities -Clear explanation of grammar rules - Phonetics and pronunciation -Dictation and spelling 			
e. Teachers' views of learners important needs in textbooks	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Appropriate topics for students' local and international needs -Appropriate texts and activities to students' level -Encourage students to develop learning strategies (through different tip-offs and directions on how to acquire the language) -Focus on students weaknesses (like auxiliary verbs and tenses) - Promote students' production not memorization (through providing more opportunities for students to use the language with partners and teacher) -Learners-centered activities 			
f. Teachers' views of their own needs in textbooks	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Can be modified and edited through the availability of soft version of materials -Whole package with supplementary materials and CDs -Teachers friendly -Clear instructions and easy to be taught - Publisher available for questions and feedback 			
g. Technology tools needed with the English materials (textbooks)	Qualitative data (brainwriting sessions)	<ul style="list-style-type: none"> -Internet-based activities -List of online sites for further learning -More Language lab activities 			

<p>h. Research and social skills are important in English language textbooks</p>	<p>Qualitative data (brainwriting sessions)</p>	<p>-Research projects -Analyzing graphs and tables -Research-based activities -Employability skills -Themes related to current affairs, health and life style</p>			
<p>i. Use of current methodologies (post-communicative approach)</p>	<p>Qualitative data (brainwriting sessions)</p>	<p>-Through the application of approaches on English as an international language such as post-communicative approach (Especially content language integrated learning (CLIL) and task- learning) .Also suggested by Ur, p.(2012)</p>			
<p>3. Institutional needs</p>					
<p>1. The Ministry of higher Education concerns when selecting and evaluating teaching materials are catered for:</p> <p>a. Appropriate for students' Proficiency levels</p> <p>b. Help students to succeed in their majors and degree programmes</p> <p>c. Align with Foundation Programmes National standards</p> <p>d. The price of the textbooks is reasonable.</p> <p>2. The materials include "innovative methodologies and groundbreaking strategies...to address the need of the new generations"</p> <p>3. The materials allow for teachers' and students' feedback.</p> <p>4. The materials can provide methods for "cross check of the students' performance using standardized international tests"</p>	<p>Assistant Director General for Academic Affairs in the Ministry of Higher Education Feedback through a short survey</p>	<p>-The aims are clearly stated in order to compare the claimed textbooks objectives with your institution's standards and the needs of different majors</p> <p>-The materials alignment with the international criteria of proficiency levels descriptions such as CEFR and ACTFL on the textbook cover</p> <p>- The materials include satisfaction surveys for both students and teachers</p> <p>-Materials include samples for "standardized international tests" for each level.</p>			

Appendix L2: Teaching materials evaluation checklist prototype 2

Teaching Materials Evaluation Checklist for English Language Programmes (TMEC for ELP)					
Text book title:					
Author:					
Level:					
Publisher /year of publication:					
1. Quick Evaluation Checklist					
The criteria	What to look for? Please tick (√) the appropriate answer	Avaiable (2points)	Not sure (1 point)	Not available (0 point)	Evaluator's Notes
1. To what extent it is Practical?	The components needed are available (coursebook, workbook)				
	Affordable price				
2. To what extent it supports learning-teaching process?	Teachers' book				
	Tests				
	CDs				
	Suitable for self-study (easy to use/ answer keys)				
3.To what extent it is context relevent to the course, learners and teachers?	Length of course				
a. Suitability for the course	Aims				
	Syllabus				
	Exams				
b. Suitablility for learners	Age				
	Level				
	Cultural background				
c. Suitablility for teachers	Clear instructions, guides and supplementary materials				
4. To what extent it appeals to learners?	Layout				
	Visuals				
	Topics				
	Total				

The textbook(s) that passed the quick evaluation can be evaluated in details using the close evaluation checklist below					
2. Close (detailed) evaluation checklist					
A. literature review					
	What to look for? (Textbook should have minimum of 1 of the specified items in this column to get 2 points in this section)	Available (2points)	Not sure (1 point)	Not available (0 point)	Evaluator's Notes
1. Second language Acquisition principles					
a. To what extent materials help the learners to develop confidence (involving them in tasks which are challenging but achievable)?	1. Activities ask students to think and which “do not provide answers all the time”				
	2. Including topics on different situations where the students are asked to provide various solutions for the same problem.				
b. To what extent materials should require and facilitate learner self-investment (through...responsibility of making decisions...making discoveries about language themselves)?	1. Activities that help the learners to make discoveries about “how the language is used” and that ask them to discuss their findings on language use.				
	2. Activities that require students to search for the answers in groups using the internet and interviewing other teachers.				
c. To what extent materials expose learners to language in authentic use (a rich and varied input which includes unplanned, semi-planned and planned discourse)?	1. Use of authentic texts are newspapers and magazines				
	2. Use of longer texts with less editing as much as possible				
	3. Use of different accents and real conversations.				
d. To what extent materials provide the learners with opportunities to use the target language to achieve communication purposes...and to develop strategic competence?	1. Content that focuses on “meaning and form”				
	2. Activities that “encourage communication between students” such as debates and making conversations				

	3. Tips and instructions on how to speak and write effectively				
	4. Activities that allow “opportunities for students to express their own meaning in their own words”				
e. To what extent materials take into account that the learners differ in affective attitudes?	1. Texts that can “amuse, excite and stimulate” students				
	2. Topics and stories that encourage students to share feelings.				
f. To what extent help the learner to develop cultural awareness?	1. Texts and topics that allow students to appreciate their own cultures and respect other cultures				
	2. Texts that discuss issues about tolerance and acceptance				
g. To what extent reflect the reality of language use?	1. Use of language as a Lingua Franca (foreign accents and non- native speakers’ conversations)				
	2. The use of pictures and videos that show how people use language in different situations.				
	Activities that encourage the “use of language outside classroom)				
h. To what extent materials cater for the different needs of learners?	1. Representing all types of life styles: rural, urban and simple lives not only focusing on “middle-class and well-educated” ones.				
	2. Providing activities that help students to discover their needs and wants and providing clarifications on how to improve them.				
i. To what extent materials help learners after course to develop “autonomous learning”?	1. Activities that teach “real life strategies, skills” e.g. portfolios, cards, journals and pictures).				
	2. Activities that show students how to become successful learners such as dictionary uses and mobile applications on language learning.				

j. To what extent materials help learners to personalize their learning?	1. Provide activities for learners involvement through asking for their 'views and opinions'.				
	2. Activities that require students to share their favourites hobbies, applications or websites.				
2. Teaching Principles pedagogical factors	What to look for? (Textbook should have minimum of 1 of the specified items in this column to get 2 points in this section)	Available (2points)	Not sure (1 point)	Not available (0 point)	Evaluator's Notes
a. To what extent the course materials can help learners to make the most effective use of previous knowledge?	1. Questions at the beginning of the text are designed to stimulate relevant knowledge and "throughout the materials".				
	2. Activities or diagrams that connect and summarize the same theme or grammar rule throughout the textbook units.				
b. To what extent the materials provide extensive L2 input?	1. Maximize the use of L2 in the classroom				
	2. Opportunities to receive input outside the classroom				
	3. Extensive reading through graded readers that come with textbooks				
	4. And training on how to use available resources				
c. To what extent the materials provide opportunities for learners' output?	1. Oral/ written tasks within the different textbooks units.				
	2. Activities that involve students in narrating a story or event then writing about it.				
3. The ELT Curriculum Design	What to look for? (Textbook should have minimum of 1 of the specified items in this column to get 2 points in this section)	Available (2points)	Not sure (1 point)	Not available (0 point)	Evaluator's Notes
a. To what extent the method or pedagogical approach of the materials is made	1. The objectives achieved by the end of the textbook				

clear to the textbooks users especially teachers?	2. The syllabus and its type (communicative, functional, etc....)				
	3.Learners' role				
	4.Teachers' role				
b. To what extent procedures and the use of pedagogical activities is well explained to the teachers?	1. The teachers' manual include detailed instructions on how to use different activities in the textbook.				
	2. The purpose of the activity is explained in the textbook				
c. To what extent procedures and techniques in giving the feedback on the activities to the learners are explained?	1. The teachers' manual provide the different techniques for giving feedback (e.g.written/oral/ visual) for the various activities.				
	3. The feedback is well connected to the educational goal of the activity.				
B.needs analysis					
1. Students' needs	What to look for? (Textbook should have minimum of 2 of the specified items in this column to get 2 points in this section)	Avaiable (2points)	Not sure (1 point)	Not available (0 point)	Evaluator's Notes
a. To what extent materials help learners to achieve their goals from learning the English language?	1. themes on "self-development" to acquire new knowledge, new language and better opportunities in future jobs				
	2. "Communicating with other cultures" through social media.				
b. To what extent materials include all language skills that are needed by learners?	Items that are specified include				
	1. Balanced focus on the 4 skills : Speaking, listening, reading and writing as well as grammar, pronunciation, and vocabulary				
	2. Providing sufficient activities on dictionary use, dictation, and spelling games				

	3. Focus on confusing areas “difficult words and topics, how to form questions				
c. To what extent materials consider learners' differences in their study habits, learning strategies and styles?	1. use of illustrative drawings, pictures, maps and infographics				
	2.use of various available sources like internet, mobile phones, newspapers and magazines				
	3. use of entertaining and educational tools like songs, short films and documentaries				
	4.Opportunities for Student-teacher interactions and for students to speak in class and be involved in teaching and presenting parts of the lessons to the class.				
2.Teachers' needs	What to look for? (Textbook should have minimum of 3 of the specified items in this column to get 2 points in this section)	Avaiable (2points)	Not sure (1 point)	Not available (0 point)	Evaluator's Notes
a. To what extent materials consider teachers' recommendations of a good teaching materials (textbooks)?	1.current and up-to-date				
	2. Diverse and interesting in their topics and content				
	3. Compatible with the assessment system and the course objectives				
	4. Able to incorporate current trends like gamification, critical thinking skills, research skills and e-learning tasks				
	5. Can be modified and edited through the availability of soft version of materials				
	6.Clear instructions and easy to be taught				

	7. Publisher available for questions and feedback				
b. To what extent materials consider the appropriate content for listening and speaking?	1. Listening for specific information				
	2. Mini-presentations				
	3. listening activities (3 minutes or longer)				
	4. Follow description from listening texts				
	5. Dialogues and turn-taking activities				
	6. Personalized speaking activities				
c. To what extent materials consider the appropriate content for the reading and writing skills?	1. Skimming and scanning skills				
	2. Reading for fun activities				
	3. Reading comprehension activities				
	4. Paraphrasing tasks				
	5. Different genres of writing (essays, letters, emails...)				
	6. low level sentence structure activities				
	7. Punctuation activities				
	8. Using connectors to join clauses				
d. To what extent materials consider the appropriate content for (vocabulary, grammar and pronunciation)	1. Vocabulary exercises				
	2. Soft copy of vocabulary list				
	3. Using idiomatic language				
	4. Using synonymous and antonyms				
	5. Grammar in context activities as well grammar rules explanations.				

	6. Phonetics and pronunciation				
	7. Dictation and spelling				
e. To what extent materials consider the use of technology tools needed in the English language materials(textbooks)?	1. Internet-based activities				
	2. List of online sites for further learning				
	3. More Language lab activities				
f. To what extent materials consider research and social skills in English language textbooks	1. Research projects				
	2. Analyzing graphs and tables				
	3. Research-based activities				
	4. Themes related to current affairs, health and life style				
3. Institutional needs	What to look for? (Textbook should have minimum of the specified items in this column to get 3 points in this section)	Avaiiable (2points)	Not sure (1 point)	Not available (0 point)	Evaluator's Notes
a. To what extent materials consider the Ministry of higher Education concerns when selecting and evaluating teaching materials?	1. Appropriate for students' Proficiency levels				
	2. Help students to succeed in their majors and degree programmes				
	3. Align with Foundation Prgrammes National standards				
	4. "cross check of the students' performance using standardized international tests"				
	5. satisfaction surveys for both students and teachers				
b. To what extent materials can provide methods for "cross check of the students' performance using standardized international tests"?	1. Provide extra standardized tests at the end of the textbook				
	2. Provide free access to international tests				
	3. Supplementary materials that explain international tests, their standards and the language inventories required from students in each skill to succeed in them				

Appendix L3: Teaching materials evaluation checklist prototype 3

Teaching Materials Evaluation Checklist for English Language Programmes (TMEC for ELP)			
Text book title:			
Author:			
Level:			
Publisher /year of publication:			
1. Quick Evaluation Checklist			
The criteria	What to look for?	Please tick (√) the appropriate answer (yes, no, not sure)	Evaluator's Notes
1. To what extent are the materials Practical?	1. The availability of main components (coursebook, workbook)		
	2. Affordable price		
	3. appropriate size		
2. To what extent do they supports learning-teaching process?	1. Teachers' book		
	2. Tests		
	3. CDs		
	4. Suitable for self-study (editions with answer keys)		
3. To what extent are they suitable for course context?	1. Length of course		
	2. Aims		
	3. Syllabus		
	4. Exams		
5. To what extent are the materials suitable to the learners' context?	1. Age		
	2. Level		
	3. Cultural background		
6. To what extent are the materials suitable to teachers' context?	1. Clear instructions, and		
	2. Detailed manuals and guides		
	3. supplementary materials		
7. To what extent do the materials appeal to learners?	1. Layout		
	2. Visuals		
	3. Topics		
	Total score		If above 80% go to next sheet
The textbook(s) that passed the quick evaluation can be evaluated in details using the close evaluation checklist in the next sheet			
2. Close (detailed) evaluation checklist			
A. Research			
	What to look for?		Evaluator's Notes

1. Second language Acquisition principles		Please tick (√) the appropriate answer (yes, no, not sure)	
a. To what extent do materials help the learners to develop confidence?	1. Activities ask students to think and which “do not provide answers all the time”		
	2. Including topics on different situations where the students are asked to provide various solutions for the same problem.		
	3. Activities that help the learners to make discoveries about “how the language is used” and that ask them to discuss their findings on language use.		
	4. Activities that require students to search for the answers in groups using the internet and interviewing other teachers.		
b. To what extent do materials expose learners to language in authentic use	1. Use of authentic texts are newspapers and magazines		
	2. Use of longer texts with less editing as much as possible		
	3. Use of different accents and real conversations.		
c. To what extent do the materials provide the learners with opportunities to use the target language to achieve communication competence?	1. Content that focuses on “meaning and form”		
	2. Activities that “encourage communication between students” such as debates and making conversations		
	3. Tips and instructions on how to speak and write effectively		
	4. Activities that allow “opportunities for students to express their own meaning in their own words”		
d. To what extent do materials take into account that learners differ in affective attitudes?	1. Texts that can “amuse, excite and stimulate” students		
	2. Topics and stories that encourage students to share feelings.		
e To what extent do materials help the learner to develop cultural awareness?	1. Texts and topics that allow students to appreciate their own cultures and respect other cultures		
	2. Texts that discuss issues about tolerance and acceptance		
f. To what extent do materials help learners	1. Provide activities for learners involvement through asking for their’ “views and opinions”.		

to personalize their learning?	2. Activities that require students to share their favourite hobbies, applications and websites.		
2. Teaching Principles	What to look for?	Please tick (√) the appropriate answer (yes, no, not sure)	Evaluator's Notes
a. To what extent do the course materials help learners to make the most effective use of previous knowledge?	1. Questions at the beginning of the text are designed to stimulate relevant knowledge and “throughout the materials”.		
	2. Activities or diagrams that connect and summarize the same theme or grammar rule throughout the textbook units.		
	3. Detailed explanation for teachers on activities that help students recall and make use of their experiences in learning		
b. To what extent do the materials help teachers to provide extensive use of second language in and outside the classroom ?	1. Maximize the use of second language in the classroom		
	2. Opportunities to receive input outside the classroom		
	3. Extensive reading through graded readers that come with textbooks		
	4. Instructions for teachers on training students on how to use available resources		
c. To what extent do the materials help teachers to provide opportunities for learners' language production?	1. Oral/ written tasks within the different textbooks units.		
	2. Activities that involve students in narrating a story or event then writing about it.		
3. The ELT Curriculum Design	What to look for?	Please tick (√) the appropriate answer (yes, no, not sure)	Evaluator's Notes
a. To what extent the methods of teaching in the materials are made clear to their users especially teachers?	1.The objectives achieved by the end of the textbook		
	2.The syllabus and its type (communicative, functional, etc....)		
	3.Learners' role		
	4.Teachers' role		
b. To what extent procedures and the use of pedagogical activities are well explained to the teachers?	1.The teachers' manual include detailed instructions on how to use different activities in the textbook.		
	2. The purpose of the activity is explained in the textbook		

c. To what extent procedures and techniques in giving the feedback on the activities to the learners are explained?	1. The teachers' manual provide the different techniques for giving feedback (e.g.written/oral/visual) for the various activities.		
	2. The feedback is well connected to the educational goal of the activity.		
B.needs analysis			
1. Students' needs	What to look for?	Please tick (√) the appropriate answer (yes, no, not sure)	Evaluator's Notes
a. To what extent do materials help learners to achieve their goals from learning the English language?	1. Topics and tasks that encourage "self-development";;		
	2. Tasks that provide guidance to acquire new knowledge		
	3. Tips on how to acquire the new language a		
	4. Thems on preparing studends for future jobs		
	2. Tasks that help students to "communicating with other cultures" through social media.		
b. To what extent do materials consider learners' differences in their study habits, learning strategies and styles?	1. use of illustrative drawings, pictures, maps and infographics		
	2.use of various available sources like internet, mobile phones, newspapers and magazines		
	3. use of entertaining and educational tools like songs, short films and documentaries		
	4.Opportunities for Student-teacher interactions and for students to speak in class and be involved in teaching and presenting parts of the lessons to the class.		
2.Teachers' needs	What to look for?	Please tick (√) the appropriate answer (yes, no, not sure)	Evaluator's Notes
a. To what extent do materials consider teachers' norms about the appropriate textbooks?	1.current and up-to-date		
	2. Diverse and interesting in their topics and content		
	3. Compatible with the assessment system and the course objectives		
	4. Able to incorporate current trends like gamification, critical thinking skills, research skills and e-leaning tasks		

	5. Can be modified and edited through the availability of soft version of materials		
	6. Clear instructions and easy to be taught		
	7. Publisher available for questions and feedback		
b. To what extent do materials consider the appropriate content for listening and speaking?	1. Listening for specific information		
	2. Mini-presentations		
	3. listening activities (3 minutes or longer)		
	4. Follow description from listening texts		
	5. Dialogues and turn-taking activities		
	6. Personalized speaking activities		
c. To what extent do materials consider the appropriate content for the reading and writing skills?	1. Skimming and scanning skills		
	2. Reading for fun activities		
	3. Reading comprehension activities		
	4. Paraphrasing tasks		
	5. Different genres of writing (essays, letters, emails...)		
	6. low level sentence structure activities		
	7. Punctuation activities		
	8. Using connectors to join clauses		
d. To what extent do materials consider the appropriate content for (vocabulary, grammar and pronunciation)	1. Vocabulary exercises		
	2. Soft copy of vocabulary list		
	3. Using idiomatic language		
	4. Using synonymous and antonyms		
	5. Grammar in context activities as well grammar rules explanations.		
	6. Phonetics and pronunciation		
	7. Dictation and spelling		
e. To what extent do materials make use of the latest technology methods ?	1. Internet-based activities		
	2. List of online sites for further learning		
	3. More Language lab activities		
f. To what extent materials incorporate research skills in their content?	1. Research projects		
	2. Analyzing graphs and tables		
	3. Research-based activities		

3. Institutional needs	What to look for?	Please tick (√) the appropriate answer (yes, no, not sure)	Evaluator's Notes
a. To what extent materials consider the Ministry of higher Education standards when selecting and evaluating teaching materials?	1. Help students to succeed in their majors and degree programmes		
	2. Align with Foundation Prgrammes National standards		
	3. satisfaction surveys for both students and teachers		
b. To what extent materials can provide methods for “cross check of the students’ performance using standardized international tests”?	1. Provide extra standardized tests at the end of the textbook		
	2. Provide free access to international tests		
Total score			

Appendix L4: Teaching materials evaluation checklist prototype 4

Teaching Materials Evaluation Checklist for English Language Programmes (TMEC <i>for</i> ELP)			
Text book title:		Clear	
Author:			
Level:			
Publisher /Year of publication:			
1. Quick Evaluation Checklist			
The criteria	What to look for?	Please select (yes, no or not sure)	Evaluator's Notes
1. To what extent are the materials Practical?	1. Availability of main components (coursebook, workbook)	▼	
	2. Affordable price	▼	
	3. appropriate size for students	▼	
2. To what extent do they supports learning-teaching process?	1. Teachers' book	▼	
	2. Tests	▼	
	3. CDs	▼	
	4. Suitable for self-study (editions with answer keys)	▼	
3. To what extent are they suitable for course context?	1. Length of course (terms /semesters)	▼	
	2. Aims	▼	
	Syllabus: Meaning here is "a list that specifies all the things that are to be taught in a course...items (words, grammatical feature, topics), or process ones (tasks) or communicative can do' (standards)." Usually evaluated through table of contents.	▼	
	4. Exams	▼	
5. To what extent are the materials suitable to the learning context?	1. Age	▼	
	2. Language Proficiency Level	▼	
	3. Cultural background	▼	
6. To what extent are the materials suitable to teaching context?	1. Clear instructions for teachers	▼	
	2. Detailed manuals and guides	▼	
	3. supplementary materials availability	▼	
7. To what extent do the materials appeal to learners?	1. Layout	▼	
	2. Visuals	▼	
	3. Topics	▼	
Total		0	If above 80% go to the detailed evaluation
$Total = \frac{\text{No. of times "Yes" selected} + \text{No. of times "No" selected}}{\text{No. of times "What to look for?" items}} \times 100$			

The textbook(s) that passed the quick evaluation can be evaluated in details using the close evaluation checklist in this sheet			
Teaching Materials Evaluation Checklist for English Language Programmes (TMEC for ELP)			
Text book title:		Clear	
Author:			
Level:			
Publisher /Year of publication:			
2. Detailed evaluation checklist			
A. Research			
1. Second language Acquisition principles	What to look for?	Please select the appropriate answer : 1: Yes 2: No 3: Not sure	Evaluator's Notes
a. To what extent do materials help the learners to develop confidence?	1. Activities that ask students to think “do not provide answers all the time”	▼	
	2. Including topics on different situations where the students are asked to provide various solutions for the same problem.	▼	
	3. Activities that help the learners to make discoveries about “how the language is used” and that ask them to discuss their findings on language use.	▼	
	4. Activities that require students to search for the answers in groups using the internet and interviewing other teachers.	▼	
b. To what extent do materials expose learners to language in authentic use (authentic means here: “Communication by and for native speakers, writers or readers in that language”)	1. Use of authentic texts as newspapers and magazines	▼	
	2. Use of longer texts with less editing as much as possible	▼	
	3. . Examples of different accents and real conversations within the textbook especially in listening and speaking	▼	
c. To what extent do the materials provide the learners with opportunities to use the target language to achieve communication competence?	1. Content that focuses on “meaning and form”	▼	
	2. Activities that “encourage communication between students” such as debates and making conversations	▼	
	3. Instructions on how to be an effective learner especially in productive skills (speaking and writing)	▼	
	4. Activities that allow “opportunities for students to express their own meaning in their own words”	▼	
d. To what extent do materials take into account that learners differ in their affective attitudes?	1. Texts that can “amuse, excite and stimulate” students	▼	
	2. Topics and stories that encourage students to share feelings.	▼	
e. To what extent do materials help the learner to develop cultural awareness?	1. Texts and topics that allow students to appreciate their own cultures and respect other cultures	▼	
	2. Texts that discuss issues about tolerance and acceptance	▼	
f. To what extent do materials help learners to personalize their learning? (adaptation to a students’ unique combination of goals, interests, and competencies and the ongoing process of shifting instruction as these conditions change)	1. Provide activities for learners’ involvement through asking for their experiences and “views and opinions”.	▼	
	2. Activities that require students to share their favourite hobbies, applications and websites.	▼	
2. Teaching Principles	What to look for?		Evaluator's Notes
a. To what extent do the course materials help learners to make the most effective use of previous knowledge?	1. Questions at the beginning of the text are designed to stimulate relevant knowledge in most units.	▼	
	2. Activities or diagrams that connect and summarize the same theme or grammar rule (summary of previous lessons)	▼	
	3. Detailed explanation for teachers on activities that help students recall and make use of their learned knowledge and experiences.	▼	
b. To what extent do the materials help teachers to provide extensive use of second language in and outside the classroom?	1. Maximize the use of second language in the classroom in most units	▼	
	2. Opportunities to receive input outside the classroom (summaries of previous lessons)	▼	
	3. Extensive reading through graded readers that come with textbooks	▼	
	4. Instructions for teachers to help students to use available resources	▼	
c. To what extent do the materials help teachers to provide opportunities for learners' language production??	1. Examples of oral/ and written tasks are available within the different textbooks units.	▼	
	2. Activities that involve students in narrating a story or event then writing about it.	▼	

3. The ELT Curriculum Design (Curriculum: "The overall plan or design for a course and how the content for a course is transformed into a blueprint for teaching and learning which enables the desired learning outcomes to be achieved")		What to look for?		Evaluator's Notes
a. To what extent the methods of teaching in the materials are made clear to their users especially teachers?	1.The objectives achieved by the end of e1.The objectives achieved by the end of each unit in the textbook are clear			
	2. The syllabus and its type (communicative, functional, etc....) are explained			
	3.Learners' role "processor, performer, initiator, problem solver or other" is specified in the materials			
	4.Teachers' role "consultant, guide and model for learning" is clarified			
b. To what extent the procedures and the use of pedagogical activities are well explained to the teachers?	1. The teachers' manual include detailed instructions on how to use different activities in the textbook.			
	2. The purpose of the activity is explained in the Teachers' book			
c. To what extent procedures and techniques in giving the feedback on the activities to the learners are explained?	1. The teachers' manual provide the different techniques for giving feedback (e.g.written/oral/ visual) for the various activities.			
	2. The feedback is well connected to the educational goal of the activity.			
B.needs analysis				
1. Students' needs		What to look for?		Evaluator's Notes
a. To what extent do materials help learners to achieve their goals from learning the English language?	1. Topics and tasks that encourage "self-development";			
	2. Tasks and tips that provide guidance to acquire the new language			
	3. Themes on preparing students for future jobs			
	4. Tasks that help students to "communicating with other cultures" through social media.			
b. To what extent do materials consider learners' differences in their study habits, learning strategies and styles?	1. use of illustrative drawings, pictures, maps and infographics			
	2.use of various available sources like internet, mobile phones, newspapers and magazines			
	3. use of entertaining and educational tools like songs, short films and documentaries			
	4. Opportunities for Student-teacher interactions and for students to speak in class and be involved in teaching and presenting parts of the lessons to the class.			
2.Teachers' needs		What to look for?		Evaluator's Notes
a. To what extent do the teaching materials consider the general qualities and preferences recommended by teachers?	1.Able to incorporate current trends like gamification, critical thinking skills, research skills and e-learning tasks			
	2. Can be modified through the availability of soft version of materials			
	3. Materials are compatible with the assessment system and the course objectives			
	4. Publisher available for questions and feedback			
	5. Diverse and interesting in their topics and content			
	6. Clear instructions and easy to be taught			
b. To what extent do the materials comply with teachers' recommended tasks and activities?	1- Phonetics and pronunciation activities			
	2-More listening activities (3 minutes or longer)			
	3-Personalized speaking activities			
	4-Graded reading activities			

	5. Skimming and scanning activities		
	6. Vocabulary and dictionary work tasks		
	Grammar in context activities		
	7. low level sentence structure activities		
	8. Using connectors to join clauses		
c. To what extent do the teaching materials comply with teachers' quality criteria?	1. Materials are organized in a logical way		
	2. Wide range of engaging topics		
	3. Content is interesting		
	4. Culturally appropriate		
	5. Easy adaptable materials		
	6. Aligned to curriculum objectives and students' needs		
	7. Easy and practical exercises for students and teachers		
d. To what extent do teaching materials consider the teachers' preferred content and skills?	1. Content that exploits CLIT (content language-integrated learning)		
	2. Utilizing E-learning: List of online sites for further learning, Soft copy of vocabulary list		
	3. Reading skills as skimming and scanning		
	4. Vocabulary and dictionary use skills		
	5. Organizing a paragraph skills as paraphrasing and summarizing		
	6. Different genres of writing (essays, letters, emails)		
	7. Research skills as analyzing graphs and tables, problem solving, problem solving and critical thinking		
	8. Topics that are relevant to students' lives		
3. Institutional needs	What to look for?		Evaluator's Notes
a. To what extent do materials consider the Ministry of higher Education standards when selecting and evaluating teaching materials?	1. Help students to succeed in their majors and degree programmes		
	2. Align with Foundation Programmes National standards		
	3. satisfaction surveys for both students and teachers		
b. To what extent materials can provide methods for "cross check of the students' performance using standardized international tests"?	1. Provide extra standardized tests at the end of the textbook		
	2. Provide free access to international tests		
	Total percentage	0	
$Total = \frac{No. of times "Yes" selected + No. of times "No" selected}{No. of times "What to look for?" items} \times 100$			

References

- AbdelWahab, M. M. (2013). Developing an English language textbook evaluative checklist. *IOSR Journal of Research & Method in Education*, 1(3), 55-70.
- Abidin, M. J. Z., Rezaee, A. A., Abdullah, H. N., & Singh, K. K. B. (2011). Learning styles and overall academic achievement in a specific educational system. *International Journal of Humanities and Social Science*, 1(10), 143-152.
- Adaskou, K., Britten, D., & Fahsi, B. (1990). Design decisions on the cultural content of a secondary English course for Mororcco. *ELT journal*, 44(1), 3-10.
- Acun, R. (2011). Curriculum Development in History Using Systems Approach. *Educational Sciences: Theory and Practice*, 11(2), 834-838.
- Ahmad, H., & Shah, S. R. (2014). EFL Textbooks: Exploring the Suitability of Textbook Contents from EFL Teachers' perspective. *Vfast Transactions on Education and Social Sciences*, 5(1), 12-20
- Akilli, G. K. (2008). Design based research vs. mixed methods: The differences and commonalities.1-10. Retrieved from https://www.researchgate.net/publication/242719481_DRAFT_Design_Based_Research_vs_Mixed_Methods_The_Differences_and_Commonalities
- Alavinia, P., & Siyadat, M. (2013). A comparative study of English textbooks used in Iranian institutes. *International Journal of Asian Social Science*, 3(1), 150-170.
- Alghamdi, A. H., & Li, L. (2013). Adapting design-based research as a research methodology in educational settings. *International Journal of Education and Research*, 1(10), 1-12.
- Alderson, J. C., & Beretta, A. (Eds.). (1992). *Evaluating second language education*. Cambridge: Cambridge University Press
- Alemi, M., & Mesbah, Z. (2012). Textbook evaluation based on the ACTFL standards: The case of Top Notch series. *Iranian EFL Journal*, 9(1), 162-171.
- Alemi, M., & Sadehvandi, N. (2012). Textbook evaluation: EFL teachers' perspectives on "Pacesetter Series". *English language teaching*, 5(7), 64-75
- Alkin, M. C. (2011). *Evaluation essentials from A to Z*. New York: Guilford Press.
- Alkin, M. C. & Christie, C. A. (2004). An evaluation theory tree. In M. C. Alkin (Ed.) *Evaluation roots: Tracing theorists' views and influences* (pp. 12-65). Thousand Oaks, CA: Sage.
- Al'Abri, K. (2011). The impact of globalization on education policy on developing countries: Oman an example. *Literacy Information and Computer Education Journal (LICEJ)*, 2(4), 491-502

- Al-Busaidi, S. and Tuzlukova, V. (2013) *General Foundation Programs in Higher Education in the Sultanate of Oman: Experiences, Challenges and Prospects*, Muscat: Mazoon Press and Publishing.
- Al-Harhi, H. K. (2002). Globalization and the necessity of educational reform in the Sultanate of Oman. In *International Conference on Secondary School Education Development*, Muscat, Oman, 22-24 December.
- Alhojailan, M. I. (2012). Thematic analysis: A critical review of its process and evaluation. *West East Journal of Social Sciences*, 1(1), 39-47.
- Al-Issa, A. S. (2005). An ideological discussion of the impact of the NNESTs' English language knowledge on ESL policy implementation "A special reference to the Omani context". *Asian EFL Journal*, 7(3), 98-112.
- Allam, A. H., & Dahlan, H. M. (2013). User experience: Challenges and opportunities. *Journal of Information Systems Research and Innovation*, 3, 28-36.
- Allwright, R. L. (1981). What do we want teaching materials for? *ELT journal*, 36(1), 5-18.
- Al-Mahrooqi, R. (2012). A student perspective on low English proficiency in Oman. *International Education Studies*, 5(6), 263-271.
- Al-Mahrooqi, R., & Al-Busaidi, S. (2010). Meeting Local Needs in Materials Writing. *Advances in Language and Literary Studies*, 1(1), 91-115.
- Al-Mamari, A. (2011). *Analyzing results of placement tests in private higher education institutions*. Unpublished report, Ministry of Higher Education, Muscat: Sultanate of Oman.
- Al-Shemeli, S. (2009) Higher education in the sultanate of Oman: planning in the context of the Globalisation. Paper presented at *the IIEP Policy forum: Tertiary Education in Small States: planning in the context of the Globalisation*. 2-3 July, 2009.
- Anderson, G. L. (2002). Reflecting on research for doctoral students in education. *Educational researcher*, 31(7), 22-25.
- Anderson, G. L., & Herr, K. (1999). The new paradigm wars: Is there room for rigorous practitioner knowledge in schools and universities? *Educational researcher*, 28(5), 12-40
- Anderson, G. L., & Herr, K. (1999). The new paradigm wars: Is there room for rigorous practitioner knowledge in schools and universities?. *Educational researcher*, 28(5), 12-40.
- Anderson, T., & Shattuck, J. (2012). Design-Based research a decade of progress in education research? *Educational Researcher*, 41(1), 16-25
- Andriessen D. (2008) Combining Design-based Research and Action Research to Test Management Solutions, in Boog B., Preece J., Slagter M. & Zeelen J. (Eds) *Towards Quality Improvement of Action Research: Developing Ethics and Standards*, pp. 125–134. Rotterdam: Sense Publishers.

- Angell, J., DuBravac, S., & Gonglewski, M. (2008). Thinking globally, acting locally: Selecting textbooks for college-level language programs. *Foreign Language Annals*, 41 (3), 562–572.
- Ansary, H., & Babaii, E. (2002). Universal characteristics of EFL/ESL textbooks: A step towards systematic textbook evaluation. *The Internet TESL Journal*, 8(2), 1-9.
- Anthony, E. M. (1963). Approach, method and technique. *English language teaching*, 17(2), 63-67.
- Aschehoug, S. H., & Boks, C. (2011). Success Criteria for Implementing Sustainability Information in Product Development. In *18th International Conference on Engineering Design (ICED)*, Aug 15-18, 2011, Copenhagen, Denmark.
- Arikan, A. (2008). Topics of reading passages in ELT course books: What do our students really read? *The Reading Matrix*, 8(2), 70- 85.
- Aronson, J. (1995). A Pragmatic View of Thematic Analysis. *The Qualitative Report*, 2(1), 1-3. Retrieved from <http://nsuworks.nova.edu/tqr/vol2/iss1/3>
- Baker, E. L., & Alkin, M. C. (1973). Formative evaluation of instructional development. *Educational Technology Research and Development*, 21(4), 389-418.
- Bakker, A. (2014). Research Questions in Design-Based Research. Freudenthal Institute for Science and Mathematics Education, Utrecht University retrieved from http://www.fi.uu.nl/en/summerschool/docs2014/design_research_michael/Research%20Questions%20in%20DesignBasedResearch2014-08-26.pdf.
- Bakker, A., & Van Eerde, H. A. A. (2015). An introduction to design-based research with an example from statistics education. In A. Bikner-Ahsbals, C. Knipping, & N. Presmeg (Eds.), *Doing qualitative research: methodology and methods in mathematics education*. New York: Springer.
- Balachandran, D. (2014) Criteria-based Post-use Evaluation of English Textbooks. *International Journal of English Language, Literature and Humanities (IJELH)*. 2(5), 72-88.
- Banegas, D. L. (2014). An investigation into CLIL-related sections of EFL coursebooks: issues of CLIL inclusion in the publishing market. *International Journal of Bilingual Education and Bilingualism*, 17(3), 345-359.
- Baporikar, N. (2012) Emerging trends in tourism industry in Oman In P. Ordóñez de Pablos, R.D. Tennyson and J. Zhao (eds) *Global Hospitality and Tourism Management Technologies* (pp. 116–135). Hershey, PA, USA: IGI Global.

- Barab, S., & Squire, K. (2004). Design-based research: Putting a stake in the ground. *The journal of the learning sciences*, 13(1), 1-14.
- Barab, S. A., Baek, E. O., Schatz, S., Scheckler, R. & Moore, J. (2008). Illuminating the braids of change in a web-supported community. In Kelly, A. E., Lesh, R. A., & Baek, J. Y. (Eds.). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*. (pp. 320-350), New York: Routledge.
- Barjesteh, H., Mukundan, J., & Vaseghi, R. (2014). A synthesis of language learning strategies: current issues, problems and claims made in learner strategy research. *Advances in Language and Literary Studies*, 5(6), 68-74.
- Bassey, M. (1995). *Creating education through research: a global perspective of educational research for the 21st century*. Newark: Kirklington Moor Press in conjunction with the British Educational Research Association
- Beyer, B. K. (1995). *How to conduct a formative evaluation*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Bichelmeyer BA. Checklist for formatting checklists [monograph on the Internet] (2003). Michigan: Western Michigan University—Evaluation Checklists Project; Available: www.wmich.edu/evalctr/checklists.
- Bickman, L., & Rog, D. J. (Eds.). (1998). *Handbook of applied social research methods*. Sage.
- Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code Development*. Thousand Oaks, CA: Sage Publications.
- Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology*, 3(2), 77-101
- Breen, M.P., & Candlin. C.N. (1987). Which materials? A consumer's and designer's guide. In L. E. Sheldon (Ed.), *ELT textbooks and materials. Problems in evaluation and development* (pp. 13-28). Oxford: Modern English Publications.
- Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *The journal of the learning sciences*, 2(2), 141-178.
- Brown, J. D. (1995). *The elements of language curriculum: A systematic approach to program development*. Boston, MA : Heinle & Heinle.
- Bülent, İ. N. A. L. (2006). Coursebook Selection Process and Some of the Most Important Criteria to be taken into Consideration in Foreign Language Teaching. *Cankaya University Journal of Arts and Sciences*, 1(5). 19-29.
- Bulger, M. (2016). Personalized Learning: The Conversations We're Not Having. Data and Society Working Paper. Available from https://datasociety.net/pubs/ecl/PersonalizedLearning_primer_2016.pdf

- Burstein, L., Freeman, H. E., Sirotnik, K. A., Delandshere, G., & Hollis, M. (1985). Data Collection: The Achilles Heel of Evaluation Research. *Sociological Methods Research*, 14 (1), 65-80.
- Byrd, P. & Celce Marcia, M. (2001), 'textbooks evaluation checklists' (Appendix B) in P. Byrd 'textbooks: evaluation for selection and analysis for implementation' in M. Celce Murcia, (ed.) *Teaching English as Second or Foreign Language* (3rd Ed.). (pp. 415-427) .Thomson Learning Inc: Heinle & Heinle.
- Byrd, P., & Schuemann, C. (2014). English as a second/foreign language textbooks: How to choose them-how to use them. In Brinton, D.M., Celce-Murcia, M., & Snow, M.A. (Eds.), *Teaching English as a second or foreign language*. (380-393). National Geographic Learning, Boston, MA: Heinle Cengage Learning
- Çakit, I. (2006). Evaluation of the EFL textbook" New Bridge to Success 3" from the perspectives of students and teachers. Unpublished MA thesis, Middle East Technical University, Ankara, Turkey.
- Cambridge Dictionaries online accessed on May 25, 2014 from <http://dictionary.cambridge.org/dictionary/british/evaluate>
- Candlin, C. N. (1984). *Syllabus design as a critical process*, *ELT Documents*, No. 118, pp. 29-46, London: Pergamon & the British Council.
- Carden, F., & Alkin, M. C. (2012). Evaluation roots: An international perspective. *Journal of MultiDisciplinary Evaluation*, 8(17), 102-118.
- Carroll, M.I, Razvi. S., & Goodliffe, T. (2009), Using Foundation Program Academic Standards as a Quality Enhancement Tool, a paper for INQAAHE 2009. Available at <http://www.oac.gov.om/qe/oqn>
- Carter, R. & Nunan, D. (2001). *Teaching English to Speakers of Other Languages*. Cambridge: CUP.
- Cassidy, S. (2004). Learning Styles: An Overview of Theories, Models, and Measures. *Educational Psychology*, 24(4), 419–444.
- Chostelidou, D. (2011). Needs-based course design: the impact of general English knowledge on the effectiveness of an ESP teaching intervention. *Procedia-Social and Behavioral Sciences*, 15, 403-409.
- Chou, P. T. M. (2010). Advantages and disadvantages of ESL course books. *The Internet TESL Journal*, 16(11). <http://iteslj.org/Articles/Chou-CourseBooks.html>
- Clarke, V., & Braun, V. (2013). Teaching thematic analysis: Overcoming challenges and developing strategies for effective learning. *The psychologist*, 26(2), 120-123. Available from: <http://eprints.uwe.ac.uk/21155>
- Coffield, F., Moseley, D., Hall, E. & Ecclestone, K. (2004) *Learning Styles and Pedagogy in Post-16 Learning. A Systematic and Critical Review*. London: Learning and Skills Research Centre.

- Collins, A., Joseph, D., & Bielaczyc, K. (2004). Design research: Theoretical and methodological issues. *The Journal of the learning sciences*, 13(1), 15-42.
- Commission of English Language Program Accreditation (CEA). (2013). Standards retrieved from <http://cea-accredit.org/about-cea/standards>
- Conner, R.F., Altman, D.G. & Jackson, C. (1984). 1984: A brave new world for evaluation. In R.F. Conner, D.G. Altman & C. Jackson (Eds.). *Evaluation Studies Review Annual* (Vol. 9). (pp. 12-22). Beverly Hills, CA: Sage.
- Cobb, P., Confrey, J., DiSessa, A., Lehrer, R., & Schauble, L. (2003). Design experiments in educational research. *Educational researcher*, 32(1), 9-13.
- Cook, V. (2008) *Second Language Learning and Language Teaching*. 4th edition, London: Hodder Education.
- Cotton, W., Lockyer, L. & Brickell, G. J. (2009). A Journey through a Design-Based Research Project. In G. Siemens & C. Fulford (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2009* (pp. 1364-1371). Chesapeake, USA: Association for the Advancement of Computing in Education.
- Crede, M., & Kuncel, N. R. (2008). Study habits, skills, and attitudes: The third pillar supporting collegiate academic performance. *Perspectives on Psychological Science*, 3(6), 425 – 453
- Creswell, J. W. (2007). *Qualitative inquiry & research design: Choosing among five approaches* (2nd ed.). Thousand Oaks, CA: Sage.
- Cronbach, Lee J. Evaluation for course improvement. *Teachers College Record* 64, no. 8 (1963): 672-683.
- Cunningsworth A (1995). *Choosing Your Coursebook*. Oxford: Heinemann Publishers Ltd.
- Daniels, A. (2011) Deputy Director for Academics, Language Center, Sultan Qaboos University (an interview with Serгон, V. 2011).
- Dede, C. (2005). Why design-based research is both important and difficult? *Educational Technology*, 45(1), 5-8.
- Denzin, N. K., & Lincoln, Y. S. (1994). The Fifth Moment. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of Qualitative Research* (pp. 575-586). Thousand Oaks, CA: Sage Publications.
- Dick, W. (1980). Formative evaluation in instructional development. *Journal of instructional development*, 3(3), 3-6.
- DiSessa, A. A., & Cobb, P. (2004). Ontological innovation and the role of theory in design experiments. *The journal of the learning sciences*, 13(1), 77-103.
- Dix, K. L. (2007). DBRIEF: A research paradigm for ICT adoption. *International Education Journal*, 8(2), 113-124.
- Dougill, J. (1987). Not so obvious. In L. E. Sheldon, (Ed.). *ELT textbooks and materials: problems in evaluation and development*. ELT Documents

- 126 (pp. 29-36). London: Modern English Publications in association with The British Council.
- Easterday, M. W., Rees Lewis, D., & Gerber, E. M. (2014). Design-based research process: Problems, phases, and applications. In *Proceedings of the International Conference of the Learning Sciences*, June 23-27, 2014, Colorado, USA (pp. 317-324).
- Eckert, C. M.; Stacey, M. K. and Clarkson, P. J. (2003). The spiral of applied research: A methodological view on integrated design research. In: *Proceedings of the 14th International Conference on Engineering Design (ICED'03)*, August 2003, Stockholm, Sweden. (pp. 19-21).
- Edelson, D. C. (2002). Design research: What we learn when we engage in design. *Journal of the Learning Sciences*, 11(1), 105-121.
- Ehrman, M. E.; Leaver, B.L. & Oxford R. L.(2003)A brief overview of individual differences in second language learning. *System* 31 (3), 313–330 www.elsevier.com/locate/system
- Ellis, R. (1997). The empirical evaluation of language teaching materials. *ELT journal*, 51(1), 36-42.
- Eri, T. (2012). The best way to conduct intervention research: methodological considerations. *Quality & Quantity*, 47 (5), 1-14.
- Ertmer, P. A., & Newby, T. J. (2013). Behaviorism, Cognitivism, Constructivism: Comparing critical features from an instructional design perspective. *Performance Improvement Quarterly*, 26(2), 43-71.
- Etikan, I, Musa, S. A., & Alkassim, R. S. (2016). Comparison of Convenience Sampling and Purposive Sampling. *American Journal of Theoretical and Applied Statistics*, 5(1), 1-4.
- Fareh, S. (2010). Challenges of teaching English in the Arab world: Why can't EFL programs deliver as expected? *Procedia-Social and Behavioral Sciences*, 2(2), 3600-3604.
- Fat'hi, J., Ghaslani, R., & Parsa, K. (2015). The Relationship between Post-method Pedagogy and Teacher Reflection: A Case of Iranian EFL Teachers. *Journal of Applied Linguistics and Language Research*, 2(4), 305-321.
- Finney, D. (2002). The ELT Curriculum: A Flexible Model for a Changing World. In J. Richards & W. Renandya (Eds.), *Methodology in Language Teaching: An Anthology of Current Practice (Approaches and Methods in Language Teaching)*, (pp. 69-79). Cambridge: Cambridge University Press
- Flagg, B.N. (1990). *Formative evaluation for educational technologies*. Hillsdale, NJ: Lawrence Erlbaum Associates
- Friedman, K. (2003). Theory construction in design research: criteria: approaches, and methods. *Design studies*, 24(6), 507-522.
- Gagne, R. M., & Briggs, L. J. (1974). *Principles of instructional design*. Oxford, England: Holt, Rinehart & Winston.

- Garinger, D. (2001). Textbook evaluation. TEFL Web Journal. Retrieved from <http://www.teflweb-j.org/v1n1/garinger.html>
- Garinger, D. (2002). Textbook selection for the ESL classroom. *Center for Applied Linguistics Digest*. Retrieved from <http://www.cal.org/resources/Digest/0210garinger.htm>
- Garvin, D.A. (1988). *Managing Quality: The Strategic and Competitive Advantage*. New York: Free Press.
- George, J. W., & Cowan, J. (1999). *A handbook of techniques for formative evaluation: Mapping the student's learning experience*. London: Psychology Press.
- Giroux, H. A. (1983). Theories of reproduction and resistance in the new sociology of education: A critical analysis. *Harvard educational review*, 53(3), 257-293.
- Given, L. M. (2008). *The Sage encyclopedia of qualitative research methods*. Los Angeles, Calif: Sage Publications.
- Goldkuhl G (2004) Meanings of Pragmatism: Ways to Conduct Information Systems Research. Presented at the *2nd International Conference on Action in Language, Organisations and Information Systems (ALOIS 2004)*. Linköping, Sweden 17-18 March 2004.
- Goldkuhl, G. (2011). Design Research in Search for a Paradigm: Pragmatism Is the Answer. In Helfert, M. & Donnellan, B (Eds.) *Practical Aspects of Design Science: European Design Science Symposium, EDSS 2011, Leixlip, Ireland, October 14, 2011, Revised Selected Papers*. (Vol. 286, pp. 84-95). Berlin, Heidelberg: Springer.
- Graddol, D. (2008). *Why global English may mean the end of 'English as a Foreign Language*. London: ULIS & British Council.
- Gray, J. (2010). *The Construction of English: Culture, Consumerism and Promotion in the ELT Global Coursebook*. London: Palgrave Macmillan.
- Gray, J. Ed. (2013). *Critical Perspectives on Language Teaching Materials*. London: Palgrave Macmillan.
- Griffee, D., & Gorsuch, G. (2016). *Evaluating Second Language Courses*. Charlotte. NC: Information Age Publishing (IAP) Inc.
- Griffiths, C. (1995). Evaluating materials for teaching English to adult speakers of other languages. *ELT Forum* 33 (3), pp. 153-168.
- Griffiths, C. (2008) Teaching/ learning method and good language learners. In Griffiths, C. (Ed.). (2008). *Lessons from good language learners* (P.255-265). Cambridge: Cambridge University Press.
- Groundwater-Smith, S., & Irwin, J. (2011). Action Research in Education and Social Work. In Markauskaite, L., Freebody, P. & Irwin, J. (Ed.). *Methodological Choice and Design: Scholarship, Policy and Practice in Social and Educational Research* (pp. 57-69). Netherlands: Springer.

- Guidy-Oulai, A. M. (2009). The development and application of a checklist for evaluating e-learning in organizations (Doctoral dissertation, Western Michigan University).
- Hall, G. (2011) *Exploring English Language Teaching: Language in Action*. London: Routledge.
- Hall, R. (2013). Mixed methods: in search of a paradigm. In T. Le, & Q .Le, (Eds.) *Conducting research in a changing and challenging world*, (pp. 71-78) New York, Nova Science Publishers.
- Hamidi, H., Montazeri, M., Alizadeh, K., & Rezaie, J. (2015). A comparative evaluation and analysis of two general English textbooks: Four Corners 1 vs. Top Notch Fundamentals A. *Theory and Practice in Language Studies*, 5(6), 1192-1199.
- 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>
- Hanafiyyeh, M. & Koosha, M. (2014). Evaluation of the EFL Textbook “Four Corners” From the Perspectives of Students. *Indian Journal of Fundamental and Applied Life Sciences*, 4 (S3), 758-777
- Harap, H. (1934). Evaluation of Courses of Study and Textbooks. *Review of Educational Research*, 4(2), 194-198.
- Harwood, N. (2010). *English language teaching materials: Theory and practice*. New York: Cambridge University Press.
- Hedge, T. (2000). *Teaching and learning in the language classroom*. Oxford: Oxford University Press.
- Henrichsen, L. E. (1983). Teacher preparation needs in TESOL: The results of an international survey. *RELC Journal*, 14(1), 18-44.
- Henson. M. (2012). Approaches to Audit and Accreditation in developing HE Systems: a comparison", 1st World Summit on Accreditation, New Delhi, India, April 2012.
- Herrington, J., McKenney, S., Reeves, T. & Oliver, R. (2007). Design-based research and doctoral students: Guidelines for preparing a dissertation proposal. In C. Montgomerie & J. Seale (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2007* (pp. 4089-4097). Chesapeake, VA: AACE.
- Heslin, P. A. (2009). Better than brainstorming? Potential contextual boundary conditions to brainwriting for idea generation in organizations. *Journal of Occupational and Organizational Psychology*, 82(1), 129-145.
- Hickcox, L.K. (1995). Learning styles: A survey of adult learning style inventory models. In R.R. Sims and S.J. Sims (Eds.). *The importance of learning styles: Understanding the implications for learning, course design, and education* (pp. 25-47). Westport, CT: Greenwood Press.

- Hogan, R.L. (2007). The historical development of program evaluation: Exploring the past and present. *Online Journal of Workforce Education & Development*, 2(4), 1-14.
- Hogue, R.J. (2013). Epistemological Foundations of Educational Design Research. In T. Bastiaens & G. Marks (Eds.), *Proceedings of E-Learn 2013--World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education* (pp. 1915-1922). Las Vegas, NV, USA: Association for the Advancement of Computing in Education (AACE)
- Hooman, Y. (2014). The evaluation of Iranian EFL textbooks from post method principles pedagogy. *Theory and Practice in Language Studies*, 4(5), 1026-1033.
- Hovorka, D.S. (2009). "Design Science Research: A call for a pragmatic Perspective," Proceedings > Proceedings of SIGPrag Workshop. *Sprouts: Working Papers on Information Systems*, 9(71). <http://sprouts.aisnet.org/9-71>
- Howatt, A. P. R. (1984). *A History of English Language Teaching*. Oxford: Oxford University Press.
- Hammersley, M. (2007) (ed). Educational Research and Evidence-based Practice. In association with the Open University. London: Sage
- Hussin, N. I. S. M., Nimehchisalem, V., & Kalajahi, S. A. R. (2015). Developing a Checklist for Evaluating the Presentation of New Vocabulary in ELT Textbooks. *International Journal of Language Education and Applied Linguistics (IJLEAL)* 02. (2015) 27–38
- Hutchinson, T. (1987). What's underneath? An interactive view of materials evaluation. In L. E. Sheldon, (Ed.). *ELT textbooks and materials: problems in evaluation and development*. ELT Documents 126 (pp. 37-44). London: Modern English Publications in association with The British Council.
- Hutchinson, T., & Waters, A. (1987). *English for specific purposes*. Cambridge: Cambridge University Press.
- Hwang, W., & Salvendy, G. (2007). What makes evaluators to find more usability problems?: a meta-analysis for individual detection rates. In Proceedings of the 12th international conference on Human-computer interaction: interaction design and usability (pp. 499-507). Springer-Verlag
- UNESCO-IBE, (2011). World Data on Education 2010/2011, Oman, (7th Edition), Geneva: UNESCO-IBE, Publication unit. Retrieved from <http://www.ibe.unesco.org/en/document/world-data-education-seventh-edition-2010-11>
- Imenda, S. (2014). Is there a conceptual difference between theoretical and conceptual frameworks? *Journal of Social Science*, 38(2), 185–195.

- Islam, A. S., & Shuchi, I. J. (2017). Deconstruction of Method-postmethod Dialectics in English Language Teaching. *Journal of Language Teaching and Research*, 8(3), 539-547.
- Jackson, P. W. (1990). *Life in classrooms*. New York: Teachers College Press.
- Jahangard, A. (2007). Evaluation of EFL materials taught at Iranian Public high schools. *The Asian EFL Journal*, 9(2), 130-150.
- James, M.E. (2013) *Educational Assessment, Evaluation and Research: The Selected Works of Mary E. James*. London: Routledge.
- Jensen, E. (2010). *Are Learning Styles a Big Hoax? What Does the Latest Science Say About Different Learners?* Brain-Based Learning. Retrieved from:
<http://www.jensenlearning.com/news/are-learning-styles-a-big-hoax-what-does-the-latest-science-say-about-different-learners/brain-based-learning>
- Johnson, R. B. (1997). Examining the validity structure of qualitative research. *Education*, 118(2), 282-292.
- Johnston, B., & Peterson, S. (1994). The program matrix: A conceptual framework for language programs. *System*, 22(1), 63-80.
- Joseph, D. (2004). The practice of design-based research: Uncovering the interplay between design, research, and the real-world context. *Educational Psychologist*, 39(4), 235-242.
- Joseph, P. B. (2011). *Cultures of Curriculum* (2nd ed). New York: Routledge.
- Johnson, R. K. (Ed.). (1989). *The second language curriculum*. Cambridge: Cambridge University Press.
- Johnson et al. (2008). A step forward: investigating expertise in materials evaluation, *ELT Journal*, 62 (2) 157–163
- Johnson, R. (Ed.). (1989). *The Second Language Curriculum* (Cambridge Applied Linguistics). Cambridge: Cambridge University Press
- Jury, T. W. (2007) *Electronic Performance Support for E-Learning Analysis and Design: A published dissertation*, Nova Southeastern University. Computing Technology in Education (MCTE, DCTE) publisher: ProQuest.
- Juuti, K., & Lavonen, J. (2006). Design-based research in science education: One step towards methodology. *NorDiNa: Nordic Studies in Science Education*, 4, 54–68.
- Karababa, Z. C., Serbes, I. & Şahin, A. F. (2010). Evaluation of the textbook Breeze in terms of the A2 level criteria determined in the European Language Portfolio. *Novitas-Royal (Research on Youth and Language)*, 4(2), 251-25
- Karamoozian, F.M. and A. Riazi, (2008). Development of a new checklist for evaluating reading comprehension textbooks. *ESP World*, 7(3): 19-39

- Kawulich, B. B. (2004). Data analysis techniques in qualitative research. *Journal of Research in Education*, 14(1), 96-113
- Kayapinar, U. (2009). Coursebook Evaluation by English Teachers. *Inonu University Journal of the Faculty of Education (INUJFE)*, 10(1), 69-78.
- Kelly, A. (2004). Design research in education: Yes, but is it methodological? *The journal of the learning sciences*, 13(1), 115-128.
- Kelly, A. E., Lesh, R. A., & Baek, J. Y. (2008). *Handbook of design research methods in education. Innovations in science, technology, engineering, and mathematics learning and teaching*, New York, NY: Routledge.
- Kelly, A. E., Baek, J. Y., Lesh, R. A., & Bannan-Ritland, B. (2008). Enabling innovations in education and systematizing their impact. In Kelly, A. E., Lesh, R. A., & Baek, J. Y. (Eds.). *Handbook of design research methods in education: Innovations in science, technology, engineering, and mathematics learning and teaching*. (pp. 3-18), New York: Routledge.
- Kemp, J., Morrison, G., & Ross, S. (1996). *Designing effective instruction*. Upper Saddle River, NJ: Prentice Hall.
- Kennedy-Clark, S. (2013). Research by Design: Design-Based Research and the Higher Degree Research student. *Journal of Learning Design*, 6(2), 26-32
- Kiely, R. (1999) *Evaluation: the ELT manager's toolkit*. ELT Management IATEFL, Special Interest Group Newsletter No.28
- Kikuchi, K. (2005). Student and teacher perceptions of learning needs: A cross analysis. *Shiken: JALT Testing & Evaluation SIG Newsletter*, 9(2), 8-20.
- Kincheloe, J. L. (2012). *Teachers as researchers (classic edition): Qualitative inquiry as a path to empowerment*. New York: Routledge.
- Kitao, K., & Kitao, S. K. (1997). Selecting and developing teaching/learning materials. *The Internet TESL Journal*, 4(4), 20-45.
- Knowlton, D. S.(2007). I Design; Therefore I Research: Revealing DBR through Personal Narrative. *Journal of Educational Technology & Society*, 10(4), 209–223.
- Kopcha, T. J., Schmidt, M. M., & McKenney, S. (2015). Editorial" special issue on educational design research (EDR) in post-secondary learning environments. *Australasian journal of educational technology*, 31(5), I-IX.
- Kormos, J (2002) Language wants of English majors in a non-native context. *System* 30 (2002) 517–542 www.elsevier.com/locate/system
- Krashen, S. D. (1982). *Principles and practice in second language acquisition*. Oxford: Pergamon.
- Kumaravadivelu, B. (1994). The Post-method Condition: (E) merging strategies for Second/foreign language teaching. *TESOL Quarterly*, 28(1), 27-48.
- Kumaravadivelu, B. (2001). Toward a postmethod pedagogy. *TESOL Quarterly*, 35 (4), 537-560

- Kumaravadivelu, B. (2006 a). TESOL Methods: Changing tracks, challenging trends. *TESOL Quarterly*, 40 (1), 59-80
- Kumaravadivelu, B. (2006 b). *Understanding language teaching: From method to postmethod*. New Jersey: Lawrence Erlbaum Associates.
- Kurniawan, L. L. (2006) *Some Insight into Textbook Evaluation: A Pragmatic Approach*. *Jurnal Pendidikan*, 12 (1), 1-13
- Kyungsim H.N. and Alexandra G. L. (2006) Language learning strategy use of ESL students in an intensive English learning context, *System* 34 (2006) 399–415 www.elsevier.com/locate/system.
- Lee, S. M. (2013). The Development of Evaluation Theories for Foreign Language Textbooks. *Journal of Pan-Pacific Association of Applied Linguistics*, 17(2), 69-89.
- Levine, T. (2002). Stability and change in curriculum evaluation. *Studies in educational evaluation*, 28(1), 1-33.
- Lim, C. J., & Lee, S. (2007). Pedagogical usability checklist for ESL/EFL e-learning websites. *Journal of Convergence Information Technology*, 2(3), 67-76.
- Lim, Y.-K., Stolterman, E., & Tenenberg, J. (2008). The Anatomy of Prototypes: Prototypes as Filters, Prototypes as Manifestations of Design Ideas. *Transactions on Computer Human Interaction*, 15(2), 1-27.
- Littlejohn, A. & Windeatt, S. (1989). Beyond language learning: Perspectives on materials design. In R. K. Johnson (Ed.). *The second language curriculum* (pp.155-175). Cambridge: Cambridge University Press.
- Littlejohn, A. (2011). The analysis of language teaching materials: inside the Trojan Horse. In Tomlinson, B. (Ed.) *Materials Development in Language Teaching*. (Pp.179-211). Cambridge: Cambridge University Press.
- Littlejohn, A. (2012). Language teaching materials and the (very) big picture. *Electronic Journal of Foreign Language Teaching*, 9(1), 283-297.
- Litz, D. R. (2005). Textbook evaluation and ELT management: A South Korean case study. *Asian EFL journal*, 48, 1-53.
- Ma, Y., & Harmon, S. W. (2009). A case study of design-based research for creating a vision prototype of a technology-based innovative learning environment. *Journal of Interactive Learning Research*, 20(1), 75-93.
- Mack, L. (2010). The philosophical underpinnings of educational research. *Polyglossia*, 19, 1-11.
- Mackenzie, N., & Knipe, S. (2006). Research dilemmas: Paradigms, methods and methodology. *Issues in educational research*, 16(2), 193-205
- Maghsoudi, N. (2016). Post Method Pedagogy: A Plausible Choice in Iran? *Studies in English Language Teaching*, 4(2), 282. 288
- Mahfoodh, M. I. H. A., & Bhanegaonkar, S. G. (2013). New Approach for Evaluating EFLM (An Eclectic Developed Checklist). *International Journal of Scientific and Research Publications*, 3(10). 1-8.

- Mahmood, K., Iqbal, M. Z., & Saeed, M. (2009). Textbook evaluation through quality indicators: The case of Pakistan. *Bulletin of Education and Research*, 31(2), 1-27.
- Mahmood, K. (2011). Standardization of Textbook Evaluation Criteria through Development of Quality Textbook Indicators. *Pakistan 12th International Convention on Quality Improvement and 2nd ANQ Regional Conference*, Lahore PIQC Institute of Quality.
- Maley, A. (2011). Squaring the Circle: Reconciling Materials as Constraint with Materials as Empowerment. In B. Tomlinson (Ed. 2nd edition), *Materials development in language teaching* (pp. 379-402). Cambridge: Cambridge University Press
- Martz, W. (2010). Validating an evaluation checklist using a mixed method design. *Evaluation and program planning*, 33(3), 215-222.
- Morrow, K. 1981. 'Principles of communicative methodology'. In K. Johnson and K. Morrow (Eds.) *Communication in the classroom*. Harlow: Longman
- Masuhara, H. (1998). What do teachers really want from coursebooks? In B. Tomlinson (ed.), *Materials Development in Language Teaching*. Cambridge: Cambridge University Press, 239-260.
- Masuhara, H. and Tomlinson, B. (2008). General English materials. In B. Tomlinson (ed.) *English language learning materials: A critical review*. (pp.17-37). London: Continuum.
- Masuhara, H., Hann, N., Yi, Y., & Tomlinson, B. (2008). Adult EFL courses. *ELT journal*, 62(3), 294.
- Matthews, A. (1985). Choosing the best available textbook. In Matthews, A., Spratt, M., and Dangerfield, L. (pp 202 206). (eds.) *At the chalkface*. London: Edward Arnold.
- Maxwell, J. A. (1992). Understanding and validity in qualitative research. *Harvard Educational Review*, 62(3), 279-300.
- Maxwell, J. A. (2012). *Qualitative research design: An interactive approach*. (3rd Ed.) Thousand Oaks. CA: Sage.
- Maxwell J. A., Mittapalli K. (2010). Realism as a stance for mixed method research. In Tashakkori A., Teddlie C. (Eds.), *Handbook of mixed methods in social & behavioral research* (2nd ed., pp. 145–168). Thousand Oaks, CA: Sage.
- McDonough, J., Shaw, C & Masuhara H. (2013). *Materials and methods in ELT: A teacher's guide* (3rd edn). London: Blackwell.
- McGrath, I. (2013) *Teaching Materials and the Roles of EFL/ESL Teachers: Practice and theory*. London: Bloomsbury.
- McGrath, I. (2002). *Materials evaluation and design for language teaching*. Edinburgh: Edinburgh University Press.
- McKay, S. (1978). Syllabuses: Structural, situational, notional. *TESOL Newsletter*, 12(5), 11.

- McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research*. London: Routledge.
- McKenney, S., & Reeves, T. C. (2013). Systematic review of design-based research progress is a little knowledge a dangerous thing? *Educational Researcher*, 42(2), 97-100.
- McKenney, S. & Reeves, T. (2014). Educational design research. In M. Spector, M. Merrill, J. Elen & M. Bishop (Eds.) *Handbook of Research on Educational Communications Technology*, (pp. 131-140). London: Springer.
- McKenney, S., & Reeves, T. (2013). Electronic performance support for curriculum materials developers: A design research project in Sub-Saharan Africa. In T. Plomp, & N. Nieveen (Eds.), *Educational design research – Part B: Illustrative cases* (pp. 533-555). Enschede, the Netherlands: SLO.
- McKenney, S., Nieveen, N., & van den Akker, J. (2006). Design research from a curriculum perspective. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp.62-90). London: Routledge.
- McKenney, S., & Reeves, T. C. (2014). Methods of evaluation and reflection in design research. *Zeitschrift für Berufs-und Wirtschaftspädagogiek*, 27, 141-153.
- Medlock, M., Terrano, M., Wixon, D. (2002) Using the RITE Method to Improve Products: A Definition and a Case Study. *Proceedings of UPA 2002*. Orlando: Usability Professionals' Association.
- Mehrdad, A.G. (2011) A subjective needs assessment of EGP students. *Procedia - Social and Behavioral Sciences* 31 (2012) 546 – 554.
- Menaker, E. S., & Coleman, S. L. (2007). Learning styles again: Where is empirical evidence? In *Proceedings of Interservice/Industry Training, Simulation, and Education Conference (IITSEC)*, Orlando, Florida, Paper (No. 7426).
- Menezes, V. (2013). Second language acquisition: reconciling theories. *Open Journal of Applied Sciences*, 3(07), 404.
- Meurant, R. C. (2010). EFL/ESL Textbook Selection in Korea and East Asia- Relevant Issues and Literature Review. In *International Conference on Ubiquitous Computing and Multimedia Applications* (pp. 89-102). Berlin, Heidelberg: Springer.
- Miekley, J. (2005). ESL textbook evaluation checklist. *The Reading Matrix*, 5(2). Available at <https://pdfs.semanticscholar.org/8767/1772c00ce96c8aecfafbf62abc964062e989.pdf>
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage.

- Miles, M. B., Huberman, A. M., & Saldaña, J. (2013). *Qualitative Data Analysis: A Methods Sourcebook* (3rd edition). Thousand Oaks, California: SAGE Publications, Inc.
- Miller, J.B. (2002). Examining the interplay between constructivism and different learning styles. *ICOTS6: The Sixth International Conference on Teaching Statistics*, Cape Town, South Africa. July 7-12, 2002. Retrieved from http://www.stat.auckland.ac.nz/~iase/publications/1/8a4_mill.pdf
- Mishan, F. (2005). *Designing authenticity into language learning materials*. Bristol: Intellect Books.
- Morse et.al. (2002). Verification strategies for establishing reliability and validity in qualitative research. *International journal of qualitative methods*, 1(2), 13-22.
- Morton, T. (2013). Critically evaluating materials for CLIL: Practitioners' Practices and Perspectives. In Gray, J. (Ed.). *Critical perspectives on language teaching materials* (pp. 111-136). Hampshire, UK: Palgrave Macmillan.
- Mukundan, J., & Ahour, T. (2010). A review of textbook evaluation checklists across four decades (1970-2008). In Tomlinson, B., Masuhara, H. (Eds.), *Research for materials development in language learning: Evidence for best practice* (pp. 336-352). London: Continuum
- Mukundan, J. (2009). *ESL textbook evaluation. A composite framework*. Köln, Germany: Lambert Academic Publishing.
- Mukundan, J., Hajimohammadi, R., & Nimehchisalem, V. (2011). *Developing an English language textbook evaluation checklist. Contemporary Issues in Education Research*, 4(6), 21-28
- 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
- Mukundan, J., & Nimehchisalem, V. (2012). Evaluative criteria of an English language textbook evaluation checklist. *Journal of Language Teaching Research*, 3(6), 1128-1134
- Mukundan, J., & Kalajahi, S. A. R. (2013). Evaluation of Malaysian English Language Teaching Textbooks. *International Journal of Education & Literacy Studies*, 1(1), 38.
- Myles, F. (2011). Second language acquisition (SLA) research: its significance for learning and teaching issues. Center for languages Linguistics & Area Studies, 18. Retrieved from <https://www.llas.ac.uk/resources/gpg/421>
- Nahrkhalaji, S. S. (2012). An evaluation of a global ELT textbook in Iran: A two-phase approach. *International Journal of Humanities and Social Science*, 2(3), 184-191.
- Nation, I.S.P. & Macalister, J. (2010) *Language Curriculum Design*. London: Routledge.

- Nielsen, J. (1994). Heuristic evaluation. In J. Nielsen & R. L. Mack (Eds.), *Usability inspection methods*, (pp. 25-62). New York: John Wiley & Sons, Inc.
- Nieveen, N. (2010). Formative evaluation in educational design research. In T. Plomp, & N. Nieveen. (Eds.). *An introduction to educational design research* (pp. 89-102). Enschede, Netherlands: SLO –Netherlands Institute for Curriculum Development.
- Nieveen, N (2007). Formative Evaluation in Educational Design Research. – An Introduction to Educational Design Research. In the proceedings of the seminar conducted at the East China Normal University, Shanghai (PR China), November 23-26. (pp. 89-102)
- Nieveen, N., & Folmer, E. (2013). Formative Evaluation in Educational Design Research. In T. Plomp & N. Nieveen (Eds.), *Educational design research – Part A: An introduction* (pp. 153–. 169) Enschede, Netherlands: SLO –Netherlands Institute for Curriculum Development.
- Nimehchisalem, V., & Mukundan, J. (2013). Usefulness of the English language teaching textbook evaluation checklist. *Pertanika Journal of Social Sciences & Humanities*, 21(2), 797-816.
- Nimehchisalem, V., & Mukundan, J. (2015). Refinement of the English language teaching textbook evaluation checklist. *Pertanika Journal of Social Sciences & Humanities*, 23(4), 761-780.
- Nunan, D. (1991). *Language teaching methodology: A textbook for teachers*. Prentice Hall.
- Nunan, D. (2001). Second language acquisition. In R. Carter & D. Nunan (Eds.), *The Cambridge Guide to Teaching English to Speakers of Other Languages* (The Cambridge Guides, pp. 87-92). Cambridge: Cambridge University Press.
- Oakley, A. (1999). Paradigm wars: some thoughts on a personal and public trajectory. *International Journal of Social Research Methodology*, 2(3), 247-254
- Oman Accreditation Council. (2008), *Plan for an Omani Higher Education Quality Management System* (The Quality Plan) Draft v4.
- Oman Academic Accreditation Authority (2011) Institutional Reports, (Salalah College). Retrieved on April, 2013 from http://www.oaaa.gov.om/institution.aspx#Inst_ReviewDwnld
- Oman Academic Accreditation Authority (2012) Institutional Reports, (Nizwa College). Retrieved on April, 2013 from http://www.oaaa.gov.om/institution.aspx#Inst_ReviewDwnld
- Oman Academic Accreditation Authority (2013) Institutional Reports, (Sur College). Retrieved on April, 2013 from http://www.oaaa.gov.om/institution.aspx#Inst_ReviewDwnld
- Oman Accreditation Council. (2008). *Oman academic standards for general foundation programs*. Oman: Ministry of Higher Education.

- Oxford, R. (2003). Language learning styles and strategies: An overview. *Learning Styles & Strategies, Oxford, Proceedings of GALA (Generative Approaches to Language Acquisition) Conference*, 1-25.
- Parker, P. E., Fleming, P. D., Beyerlein, S., Apple, D., & Krumsieg, K. (2001). Differentiating assessment from evaluation as continuous improvement tools [for engineering education]. *31st ASEE/IEEE Frontiers in Education Conference*, 1462. pp. T3A-1. IEEE.
- Parlett, M. (1981). Illuminative evaluation. In P. Reason & J. Rowan (Eds), *Human Enquiry* (pp. 219–226). Chichester: Wiley
- Pratt, D. D. & Associates (1998). *Five Perspectives on Teaching in Adult and Higher Education*. Malabar, FL: Krieger Publishing.
- Patton, M.Q. (1975). *Alternative Evaluation Research Paradigm*. University of North Dakota Study Group on Evaluation Monograph Series, Grand Forks, North Dakota.
- Patton, M. (1990). *Qualitative evaluation and research methods*. Beverly Hills, CA: Sage.
- Payne, D.A. (1994). *Designing educational project and program evaluations: A practical overview based on research and experience*. Boston: Kluwer Academic Publishers.
- Piccardo, E. et al. (2011) *Pathways through assessing, learning and teaching in the CEFR, European Centre for Modern Languages*. Austria: Council of Europe Publishing
- Petrie, H., & Bevan, N. (2009). The evaluation of accessibility, usability and user experience. In: Stepanidis, C. (ed.). *The Universal Access Handbook*, (pp. 10–20). New York: CRC Press.
- Petrina, S. (2004). The politics of curriculum and instructional design/theory/form: Critical problems, projects, units, and modules. *Interchange*, 35(1), 81-126.
- Plomp, T., & Nieveen, N. (2007). An introduction to educational design research. In *Proceedings of the Seminar Conducted at the East China Normal University* November 23-26, Shanghai: (PR China) SLO-Netherlands Institute for Curriculum Development. Retrieved from http://www.slo.nl/downloads/2009/Introduction_20to_20education_20design_20research.pdf
- Plomp, T (2013). *Educational Design Research: An Introduction*. In *Proceedings of the Seminar Conducted at the East China Normal University*, Shanghai (PR china), November 23-26, Enschede: Netherlands Institute for Curriculum Development.
- Prabhu, N.S. (1987) *Second language pedagogy*. Oxford: Oxford University Press.
- Rahimpour, M. and Hashemi, R. (2011). Textbook selection and evaluation in EFL context. *World Journal of Education*, 1(2): 62-68.

- Rahimpour, S. (2013). TEFL textbook evaluation. Proceedings from *the Global Summit on Education*, Kuala Lumpur, Malaysia), 11-12 March. Retrieved from http://worldconferences.net/proceedings/gse2013/papers_gse2013/203%20Sepideh%20Rahimpour-2.pdf
- Rashidi, N., & Safari, F. (2011). A model for EFL materials development within the framework of critical pedagogy (CP). *English Language Teaching*, 4(2), 250.
- Ranalli, J. C. (2002). *An evaluation of New Headway Upper-Intermediate*. University of Birmingham. Retrieved from <http://www.cels.bham.ac.uk/resources/essays/Ranalli3.pdf>
- Razmjoo, S. A. (2012). Developing a textbook evaluation scheme for the expanding circle. *Iranian Journal of Applied Language Studies*, 2(1), 121-136.
- Razvi, S., Goodliffe, T., Al Habsi, F and Al Hassnawi, A. (2011) Towards an Effective Institutional Accreditation System: Lessons Learnt from the Experience of the Sultanate of Oman, paper presented at *the HEIC Conference*, Beirut, 31 October – 2 November.
- Razvi, S., Trevor-Roper, S., Goodliffe, T., Al-Habsi, F., & Al-Rawahi, A. (2012). Evolution of OAAA strategic planning: using ADRI as an analytical tool to review its activities and strategic planning. In Proceedings of *Seventh Annual International Conference on Strategic Planning for Quality Assurance and Accreditation of Universities and Educational Arab Institutions*, Cairo.
- Reeves, T. C. (2011). Can educational research be both rigorous and relevant? *Educational Designer*, 1(4), 1-24.
- Reid, J. (1987). The Learning Style Preferences of ESL Students. *TESOL Quarterly*, 1(21), 87-111.
- Reigeluth, C. M., & Carr-Chellman, A. A. (2009). Understanding instructional theory. In C. M. Reigeluth & A. A. Carr-Chellman (Eds.), *Instructional-design theories and models, volume III: Building a common knowledge base* (pp. 3–26). New York, NY: Routledge.
- Reigeluth, C.M., Beatty, B.J., & Myers, R.D. (Eds.). (2017). *Instructional-design theories and models, Volume IV: The learner-centered paradigm of education*. New York, NY: Routledge.
- Reimann, P. (2013). Design-based research: Designing as research. In R. Luckin, S. Puntambekar, P. Goodyear, B. Grabowski, J. Underwood, J., & N. Winters (Eds.) *Handbook of Design in Educational Technology*, (pp. 44–52). New York, NY: Routledge.
- Reimann, P. (2011). Design-based research. In Markauskaite, L., Freebody, P., & Irwin, J. (Eds.). *Methodological choice and design: Scholarship, policy and practice in social and educational research* (Vol. 9). Springer Science & Business Media (pp. 37-50). Netherlands: Springer.

- Rhoten, D. (2000). Making meaning of globalization for education. Paper presented in *the 44th annual meeting of the Comparative and International Education Society*. San Antonio, Texas. March 8-11.
- Riasati, M. J., & Zare, P. (2010). Textbook Evaluation: EFL Teachers' Perspectives on "New Interchange". *Studies in Literature and Language*, 1(8), 54-60.
- Richey, R. C., Klein, J. D., & Nelson, W. A. (2004). Developmental research: Studies of instructional design and development. In Jonassen D. H. (Ed) *Handbook of Research on Educational Communications and Technology*. (pp. 1099-1130) New Jersey: Lawrence Erlbaum Associates.
- Richey, R.C. & Klein, J.D. (2007). *Design and development research*. USA, New Jersey: Lawrence Erlbaum Associates, Inc. Publishers
- Richards, J. C. (1993). Beyond the text book: The role of commercial materials in language teaching. *RELC journal*, 24(1), 1-14.
- Richards, J.C, (2001) *Curriculum Development in Language Teaching*. Cambridge: Cambridge University Press.
- Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge: Cambridge university press.
- Richards, J.C. (2002) Theories of teaching in language teaching. In J.C. Richards and W.A. Renandya (eds) *Methodology in Language Teaching. An Anthology of Current Practice* (pp. 19–25). Cambridge: Cambridge university press.
- Richards, J. C. (2006). Materials development and research—Making the connection. *RELC Journal*, 37(1), 5–26.
- Richards, J. C. (2013). Curriculum approaches in language teaching: Forward, central, and backward design. *RELC Journal*, 44(1), 5-33.
- Richards, J. C. (2001). *The Role of Textbooks in a Language Program*. Cambridge, UK: Cambridge University Press.
- Richardson, V. (1994). Conducting research on practice. *Educational researcher*, 23(5), 5-10.
- Roberts, A. D. (1980). The Roberts' Checklist-Selecting and Evaluating Social Studies Materials. *The Social Studies*, 71(3), 114-117.
- Roblin, N. P., Schunn, C., & McKenney, S. (2018). What are critical features of science curriculum materials that impact student and teacher outcomes?. *Science education*, 102(2), 260-282.
- Roberts, J. T. (1996). Demystifying materials evaluation. *System*, 24(3), 375-389.
- Rubdy, R. (2003): Selection of Materials, in: Brian Tomlinson, (Ed.).Tomlinson, B. *Developing Materials for language Teaching*. (pp.43-54).London: Continuum.

- Rohrer, D., & Pashler, H. (2012). Learning Styles: Where's the Evidence?. Online Submission, 46(7), 634-635
- Roschelle, J., Bakia, M., Toyama, Y., & Patton, C. (2011). Eight issues for learning scientists about education and the economy. *Journal of the Learning Sciences*, 20(1), 3-49.
- Russell, K. (2002) Why the Culture of Academic Rigour Matters to Design Research: Or, Putting Your Foot into the Same Mouth Twice. Working Papers in Art and Design 2. Accessed from https://www.herts.ac.uk/__data/assets/pdf_file/0007/12310/WPIAAD_vol2_russell.pdf
- Sahasrabudhe, S., Murthy, S., & Iyer, S. (2013). Applying Design Based Research to create Instructional Design templates for Learning Objects. In *New Frontiers in Education* (Jan-Mar 2013)
- Sandoval, W. A., & Bell, P. (2004). Design-based research methods for studying learning in context: Introduction. *Educational Psychologist*, 39(4), 199-201
- Sandoval, W.A., (2014) Conjecture Mapping: An Approach to Systematic Educational Design Research. *Journal of the Learning Sciences*, 23:1, 18-36.
- Sari, E., & Lim, C. P. (2012). Design-based research: Understanding its application in a teacher professional development study in Indonesia. *The Asia-Pacific Education Researcher*, 21(1), 28-38.
- Seedhouse, P. (1995). Needs analysis and the general English classroom. *ELT Journal*, 49(1), 59-65.
- Schofield, J. W. (2002). Increasing the generalizability of qualitative research. In A. M. Huberman, & M. B. Miles (Eds.). *The Qualitative Researcher's Companion* (pp. 171-203). Thousand Oaks, CA: Sage Publications.
- Scholtz, J. (2004). *Usability evaluation*. Gaithersburg : IAD National Institute of Standards and Technology Encyclopedia of Human- Computer Interaction.
- Schroeter, D. C. (2008). Sustainability evaluation: Development and validation of an evaluation checklist (Doctoral dissertation, Western Michigan University).
- Scriven, M. (1991). *Evaluation thesaurus*, 4th ed. Thousand Oaks, CA: Sage Publications
- Scriven, M. (1996). The theory behind practical evaluation. *Evaluation*, 2(4), 393-404.
- Scriven, M. (2007). *The logic and methodology of checklists*. Kalamazoo, MI: The Evaluation Center. Available from <http://www.wmich.edu/evalctr/checklists/>
- Sergon, V.(2011). Playing the Blame Game: English Education in Omani Government Schools. *Independent Study Project (ISP) Collection*. Paper 1132. 1-36

- Shah, S. K., Rafique, S., Shakir, A., & Zahid, S. (2014). Textbook evaluation of English for academic purposes by British Council. *Research on Humanities and Social Sciences*, 4(7), 104-114.
- Shah, J. K., Ensminger, D. C., & Thier, K. (2015). The Time for Design-Based Research Is "Right" and "Right Now". *Mid-Western Educational Researcher*, 27(2), 152-171.
- Shariffudin, R. S. (2007). Design of instructional materials for teaching and learning purposes: theory into practice. *Malaysian Education Deans' Council Journal*, (1), (pp.97-110)
- 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. In *The First Conference on Language Learning and Teaching: An Interdisciplinary Approach*. October, 30 and 31.
- Shavelson, R. J., D. C. Phillips, L. T., Towne, L., & Feuer, M. J. (2003). On the science of education design studies. *Educational Researcher*, 32(1), 25–28.
- Sheldon, L.E. (ed) (1987) *ELT Textbooks and Materials: Problems in Evaluation and Development*, ELT Document 126, Oxford: Modern English Publications and The British Council.
- Sheldon, L (1987). Textbooks and Materials: Problems in Evaluation and Development. In Sheldon, L.E. (ed) (1987) *ELT Textbooks and Materials: Problems in Evaluation and Development*, ELT Document 126. (pp 10-18). Oxford: Modern English Publications and the British Council.
- Sheldon, L. E. (1988). Evaluating ELT textbooks and materials. *ELT Journal* 42.4, 237–246.
- Skierso, A. (1991). Textbook selection and evaluation. In M. Celce-Murcia (Ed.), *Teaching English as a second or foreign language* (pp. 432-453). Boston, MA: Heinle & Heinle.
- Smith, K. M., & Boling, E. (2009). What Do We Make of Design? Design as a Concept in Educational Technology. *Educational Technology*, 49(4), 3-17.
- Sobh, R., & Perry, C. (2006). Research design and data analysis in realism research. *European Journal of marketing*, 40(11/12), 1194-1209.
- Spector, M. J., Merrill, M. D., Elen, J., & Bishop, M. J. (Eds.). (2013). *Handbook of Research on Educational Communications and Technology*. New York: Springer Science and Business Media.
- Stenhouse, L. (1975). *An introduction to curriculum research and development* (Vol. 46). London: Heinemann.
- Stern, H. H. (1989). Seeing the wood AND the trees: Some thoughts on language teaching analysis. In Johnson, R. K. (Ed.). (1989) *The second language curriculum*. Cambridge: Cambridge University Press. Pp.207-221

- Stetler, C. B., Legro, M. W., Wallace, C. M., Bowman, C., Guihan, M., Hagedorn, H & Smith, J. L. (2006). The role of formative evaluation in implementation research and the Queri experience. *Journal of General Internal Medicine*, 21(2), S1-S8.
- Stokes, D. E. (1997). *Pasteur's quadrant: Basic science and technological innovation*. Washington, DC: Brookings Institution Press.
- Stufflebeam, D. L. (1987). The CIPP Model for Program Evaluation. In Madus, G.F. et al (Eds). *Evaluation Models: Viewpoints on Educational and Human Services*. Boston: Kluwer Nijhoff Publishing.
- Stufflebeam, D. L. (2000). *Guidelines for developing evaluation checklists: The checklists development checklist* (CDC).Kalamazoo, MI: Center for Evaluation, Western Michigan University.
- Stufflebeam, D. L. (2001). Evaluation models. In G. T. Gary & J. C. Greene (Eds.), *New Directions for Evaluation* (Vol. 89, pp. 7-99). San Francisco: Jossey-Bass.
- Stufflebeam, D. L. (2007). CIPP Evaluation model checklist: A tool for applying the CIPP Model to assess long-term enterprises. *Evaluation Checklists Project*. Available from [https://wpqr4.adb.org/LotusQuickr/cop-mfdr/PageLibrary482574A2002211AD.nsf/h_Toc/F342DF5BFF43539048257516002F1FC4/\\$FILE/Models%20for%20program%20evaluation_Stufflebeam%202001.pdf](https://wpqr4.adb.org/LotusQuickr/cop-mfdr/PageLibrary482574A2002211AD.nsf/h_Toc/F342DF5BFF43539048257516002F1FC4/$FILE/Models%20for%20program%20evaluation_Stufflebeam%202001.pdf)
- Stufflebeam, D. L., & Shinkfield, A. J. (2012). *Systematic evaluation: A self-instructional guide to theory and practice*. Boston, MA: Kluwer-Nijhoff Publishing.
- Symonds, J. E., & Gorard, S. (2008). The death of mixed methods: research labels and their casualties. In *BERA Annual Conference*, Heriot Watt University, Edinburgh. Jossey-Bass.
- Su, S. W. (2012). The Various Concepts of Curriculum and the Factors Involved in Curricula-making. *Journal of Language Teaching and Research*, 3(1), 153-158.
- Tarone, E., & Yule, G. (1989). *Focus on the language learner: Approaches to identifying and meeting the needs of second language learners*. Oxford: Oxford University Press.
- Tasnimi, M. (2014). The Role of Teacher in the Postmethod Era. Express. *International Journal of Multi-Disciplinary Research*, 1(3), 1-8.
- Tashakkori, A., Teddlie, C. (2003). *Handbook of mixed methods in social and behavioral research*. Thousand Oaks, CA: Sage.
- Taylor, P. C., & Medina, M. N. D. (2013). Educational research paradigms: From positivism to multiparadigmatic. *The journal of Meaning-centered education*, 1(2), 1-13. Available from <http://www.meaningcentered.org/educational-research-paradigms-from-positivism-to-multiparadigmatic/>
- Teddlie, C., & Yu, F. (2007). Mixed methods sampling a typology with examples. *Journal of mixed methods research*, 1(1), 77-100.

- Tennyson, R. D. (2010). Historical reflection on learning theories and instructional design. *Contemporary Educational Technology*, 1(1), 1-16.
- Tessmer, M. (1993). *Planning and conducting formative evaluations: Improving the quality of education and training*. London & New York: Taylor & Francis Group, Routledge.
- Tomlinson, B. (Ed.) (1998). *Materials development in language teaching*. Cambridge: Cambridge University Press.
- Tomlinson, B. (Ed.) (2003). *Developing Materials for Language Teaching*. London: Continuum.
- Tomlinson, B. Ed. (2008). *English Language Learning Materials: A Critical Review*. London: Continuum.
- Tomlinson, B., & Masuhara, H. (Eds.) (2010). *Research in Materials Development in Language Learning: Evidence for Best Practice*. London, UK: Bloomsbury.
- Tomlinson, B. (2012). Materials development for language learning and teaching. (State-of-the-Art Article). *Language Teaching* 45(2):143-179.
- Tomlinson, B. (Ed.). (2013a). *Developing Materials for Language Teaching*. London: Bloomsbury Publishing.
- Tomlinson, B. (Ed.). (2013b). *Applied Linguistics and Materials Development*. London: Bloomsbury Publishing.
- Tomlinson, B. & Masuhara, H. (2017). *A Complete Guide to the Theory and Practice of Materials Development for Language Learning*. Chichester: Wiley.
- Tomlinson, B., & Masuhara, H. (2013). Adult coursebooks. *ELT journal*, 67(2), 233-249.
- Tomlinson, B., Dat, B., Masuhara, H., & Rubdy, R. (2001). EFL courses for adults. *ELT journal*, 55(1), 80-101.
- Tsiplakides, I. (2011). Selecting an English coursebook: theory and practice. *Theory and Practice in Language Studies*, 1(7), 758-764.
- Tucker, C. A. (1975). Evaluating beginning textbooks. *English Teaching Forum*, 13, 355-361
- Tucker, C.A. (1975). Evaluating beginning Textbooks, *English Teaching forum* 13: 3/ 4 (Special issue Pt 2); 355-361 [Originally printed in 1968 in *English Teaching Forum*, 6, 5, pp.8-15, and subsequently reprinted as Appendix 3 in H. Madson and J. Brown (1978), *Adaptation in Language Teaching*, Rowley, MA: Newbury House, 219-237].
- Tyler, R. W. (2013). *Basic principles of curriculum and instruction*. Chicago: the University of Chicago press.
- Ur, P. (1996). *A course in language teaching: Practice & Theory* (pp. 184-187). Cambridge: Cambridge University Press.

- Van den Akker, J. (1999). Principles and methods of development research. In Jvan den Akker,J , Branch, R. M., Gustafson, K., Nieveen, N. & Plomp, T. (Eds.). *Design Approaches and Tools in Education and Training* (pp. 1–14). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Van den Akker, J., Gravemeijer, K., McKenney, S., & Nieveen, N. (Eds.). (2006). *Educational design research*. London: Routledge.
- Van den Akker, J. (2007) Curriculum design research. An introduction to educational design research- in the proceedings of *the seminar conducted at the East China Normal University, Shanghai* (PR China), November 23-26, 2007, 37-52
- Van den Akker, J. (2010). Building Bridges: How Research May Improve Curriculum Policies and Classroom Practices. In S. M. Stoney (Ed.), *Beyond Lisbon 2010: Perspectives from Research and Development for Educational Policy in Europe* (pp. 175-195). Sint-Katelijne-Waver, Belgium: CIDREE.
- VanGundy, A. B. (1984). Brain writing for new product ideas: an alternative to brainstorming. *Journal of Consumer Marketing*, 1(2), 67-74.
- Vasalou, A. (2015) A Critical Comparison of Design-Based Research and Action Research. UK, London Knowledge Lab.
- Vesper, J., Reeves, T., & Herrington, J. (2011). The application of expert review as a formative evaluation strategy within an educational design research study.
- Vesper, J., Reeves, T.C. & Herrington, J. (2011). The application of expert review as a formative evaluation strategy within an educational design research study. In *Proceedings of World Conference on E-Learning in Corporate, Government, Healthcare, and Higher Education 2011* (pp. 973-978). Chesapeake, VA: AACE.
- Visscher-Voerman, I., Gustafson, K., & Plomp, T. (1999). Educational design and development: An overview of paradigms. In Jvan den Akker,J , Branch, R. M., Gustafson, K., Nieveen, N. & Plomp, T. (Eds.). *Design Approaches and Tools in Education and Training* (pp. 15-28). Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Walker, D. (2006). Towards productive design studies. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational design research* (pp. 9–18). London: Routledge.
- Walker, R. (2011). Design-Based Researchresearch Design-based (DBR): Reflections on Some Epistemological Issues and Practices. In Markauskaite, L., Freebody, P. & Irwin, J. (Ed.). *Methodological Choice and Design: Scholarship, Policy and Practice in Social and Educational Research* (pp. 51-56). Netherlands: Springer.
- Walker-Egea, C. F. (2014). Design and Validation of an Evaluation Checklist for Organizational Readiness for Evaluation Capacity Development (Doctoral dissertation, University of South Florida).

- Wang, Wenjuan and Duffy, Alex H.B. (2009) A triangulation approach for design research. In: *Proceedings of International Conference on Engineering Design, ICED'09*. The Design Society, pp. 275-286. ISBN 9-781904-670063
- Wang, F., & Hannafin, M. J. (2005). Design-based research and technology-enhanced learning environments. *Educational technology research and development*, 53(4), 5-23.
- Wen-Cheng, W., Chien-Hung, L., & Chung-Chieh, L. (2011). Thinking of the Textbook in the ESL/EFL Classroom. *English Language Teaching*, 4 (2), 91-96
- Weir, C. J., & Roberts, J. (1994). *Evaluation in ELT*. Oxford, UK: Blackwell.
- Weng, P.(2012)The Effect of Learning Styles on Learning Strategy Use by EFL Learners, *Journal of Social Sciences* 8 (2): 230-234, 2012 © 2012 Science Publications
- West, R. (1994). Needs analysis in language teaching. *Language teaching*, 27(01), 1-19.
- Weston, C., McAlpine, L., & Bordonaro, T. (1995). A model for understanding formative evaluation in instructional design. *Educational Technology Research and Development*, 43(3), 29-48.
- Williams, M., & Burden, R. (1994). The Role of Evaluation in ELT Project Design. *ELT Journal*, 48(1), 22.
- White. R. V (1988). *The ELT Curriculum: Design, Innovation and Management*. London: Basil Blackwell.
- White, A. (2001). Evaluation of a ELT Coursebook Based on Criteria Designed by McDonough and Shaw. Retrieved from <http://www.birmingham.ac.uk/Documents/college>.
- Williams, D. (1983). Developing criteria for textbook evaluation. *ELT journal*, 37(3), 251-255.
- Willis, J. W. (2007). *Foundations of qualitative research - interpretive and critical approaches*. CA: Sage Publications, Inc.
- Wilson, C. (2012). Method 22 of 100: Affinity Diagramming. Available from <http://dux.typepad.com/files/Method%2022%20of%20100.pdf>
- Wilson, C. (2013). *Brainstorming and beyond: a user-centered design method*. Newnes. Chicago.
- Wilson, B. G. (2014). Knowledge creation in design-based research projects: Complementary efforts of academics and practitioners. In T. Kopcha, C. Buston, & K. Moore (Eds.), *Proceedings of the Design-based Research Conference*. Retrieved from: <http://dbrxroads.coe.uga.edu/proceedings.html>.
- Wilson, C. (2013). *Credible checklists and quality questionnaires: A user-centered design method*. Waltham, MA: Morgan Kaufmann.
- Wilson, C. (2013). *Brainstorming and Beyond A User-Centered Design Method*, UK, Oxford: Elsevier Inc.

- Wingate L.A. (2002). *The Evaluation Checklist Project: The inside Scoop on Content, Process, Policies, Impact, and Challenge*, The Evaluation Center. Western Michigan University. Accessed from <https://wmich.edu/sites/default/files/attachments/u350/2014/insidescoop.pdf>
- Wisniewska, H. (2011). Modern teaching materials: SWOT analysis of an ESP textbook. *International conference, the future of Education* (p. 1). Florence: Pixel.
- Worthen, B. R., & Sanders, J. R. (1973). Evaluation as disciplined inquiry. In B. Worthen & J. Sanders (Eds.), *Educational evaluation: Theory and practice* (pp. 10-39). Worthington, OH: Charles A. Jones.
- Yin, R. K. (2003). *Case study research: Design and methods* (3rd Ed.). Thousand Oaks, CA: Sage.
- Youker, B. W., & Ingraham, A. (2014). Goal-Free Evaluation: An Orientation for Foundations' Evaluations. *The Foundation Review*, 5(4), Article 7.
- Zachariadis, M., Scott, S. V., & Barrett, M. I. (2013). Methodological Implications of Critical Realism for Mixed-Methods Research. *MIS quarterly*, 37(3), 855-879.
- Zahan, I., & Begum, J. (2013). How to Evaluate an EFL/ESL Textbook-a Problem and a Solution. *ASA University Review*, 7(1), 193-202.
- Zohrabi, M. (2011). Coursebook Development and Evaluation for English for General Purposes Course. *English Language Teaching*, 4(2), 213-222.
- Zhu, W., & Liao, F. (2008). On Differences between General English Teaching and Business English Teaching. *English Language Teaching*, 1(2), 90-95.