



## Monitoring “Testing Language Skills: From Theory to Practice” from the Perspective of Assessment Literacy

Rahman Sahragard

Associate professor in Applied Linguistics at Shiraz University

rsahragard@rose.shirazu.ac.ir

Fatemeh Zolfaghari (Corresponding author)

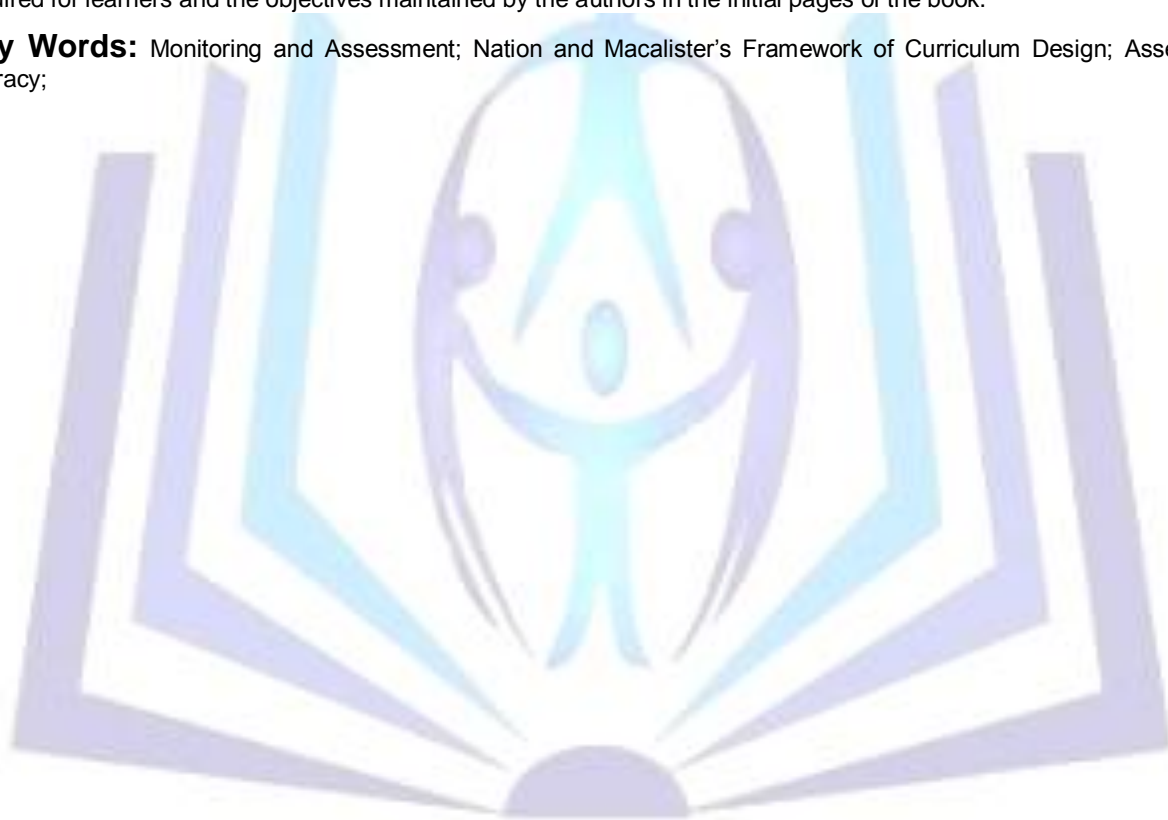
PhD student of TEFL at Shiraz University

fatemezolfaghari@ymail.com

### ABSTRACT

The present study was conducted in order to overcome the problems associated to the unsatisfied assessment literacy needs of the university students with the aid of textbook monitoring procedure. To this end, using the teachers' recommendations, the book entitled “testing language skills: from theory to practice” which is a well-known book being taught at Iranian BA testing courses of English translation and English literature was analyzed from the perspective of assessment literacy and in line with the monitoring and assessment dimension of Nation and Macalister's (2010) framework of curriculum design. This analysis was instantiated with interviews regarding learners' assessment literacy expectations of their BA testing courses. Then monitoring was performed based on the identified components that were required for learners and the objectives maintained by the authors in the initial pages of the book.

**Key Words:** Monitoring and Assessment; Nation and Macalister's Framework of Curriculum Design; Assessment Literacy;



## Council for Innovative Research

Peer Review Research Publishing System

**Journal:** INTERNATIONAL JOURNAL OF RESEARCH IN EDUCATION METHODOLOGY

Vol. 6, No.2

[www.ijrem.com](http://www.ijrem.com) , [ijremeditor@gmail.com](mailto:ijremeditor@gmail.com)



## INTRODUCTION

The demands of the interpretive epistemology and the need for evaluative procedures that should be suited to the requirements of each specific context call for promoting present learners' - as future teachers' - awareness of assessment and monitoring procedures (Fulcher, 2012; Popham, 2009). This is because knowledge is socially constructed, and assessment of the knowledge is in turn connected to the social context of its application and thus, no prescription is acceptable for all contexts (Fulcher, 2012).

Apart from the notion of context, teachers' professional development is strongly correlated with their levels of assessment knowledge (Popham, 2009). It should be noted that materials development practices can lead to desired levels of teachers' professional development, as well (Nunez Pardo, & Tellez Tellez, 2009). Additionally, Webb (2002) maintains that the importance of assessment literacy is due to two main reasons. These reasons include the emergence of standard-based movements that have raised students' explicitly described expectations, thereby increasing the need for measuring the extent to which learners' expectations are satisfied (Webb, 2002). The other reason is employing various agreed upon forms of assessment in today's academic situations (Webb, 2002).

Within the broad scope of curriculum design, textbooks play a pivotal role. This role has not been limited to provision of the course content, but rather, if not all, to most of the dimensions of curriculum evaluation. Thus, it is implied that any attempt to monitor and assess the course curriculum can narrow down its focus on the textbooks around which the course revolves. Taking the monitoring and assessment dimension of Nation and Macalister's (2010) framework for curriculum development, the present study aimed at monitoring and assessing the book developed by Farhady, Jafarpour and Birjandi (2007) as one of the main materials applied in BA testing courses at Iranian universities.

Due to the wide spectrum of stakeholders and materials that are involved in the process of curriculum design and the beneficial roles of textbooks, the scope of the present enquiry was narrowed down to textbooks and learners' expectations of the textbooks in order to achieve detailed understanding of the process of assessment literacy. Accordingly, this study was tuned to consider assessment literacy by virtue of Nation and Macalister's (2010) model of curriculum design. Considering assessment and monitoring dimension of this model, the existing positive or negative features of Farhady et al's (2007) testing course book were addressed. This aim could not be satisfied unless the aforementioned course book was analyzed in terms of the degree to which it would prepare learners for the process of assessment. Hence, the following research questions were answered on the basis of the findings of the current study:

1. Which components of assessment literacy are required for Iranian BA university students studying English translation and English literature?
2. Does the book *testing language skills: from theory to practice* satisfy the assessment literacy requirements of Iranian BA university students studying English translation and English literature?
3. To what extent the objectives of the authors of the aforementioned book have been met in its context of application?

## LITERATURE REVIEW

Most, if not all, educational contexts rely on textbooks for satisfying knowledge requirements of the learners. This is because textbook is one of the rich sources for promoting learning especially foreign language learning (Widodo, 2007, Atkinson, 2008). Clearly, the choice and application of textbooks is influenced by the various policies in the learning contexts (Mahmood, 2010). Accordingly, in some learning environments the use of multiple textbooks is the norm, whereas it is not the case with other contexts. Furthermore, textbooks reflect the cultural demands and links of every educational context (Sahragard, Rahimi, & Zaremoayehdi, 2009). Bearing all these issues in mind, assessment and monitoring of the available textbooks for resolving the existing problems facilitates development of new learning materials (Lee, 2013).

The interpretive epistemology and the socio-cultural theory as its sibling, widely acknowledge all the possible and alternative interpretations and meanings (either subjective or objective) that can be derived from an assessment (Scarino, 2013). This implies that not only test developers, but also all the people including test users, test takers, administrators, materials developers, and even learners' parents are directly involved in and connected to the assessment process (Scarino, 2013). That being so, scholars including those specialized in the area of materials development have concentrated their attention on assessment and monitoring as the interdisciplinary concepts that should be considered in development of various materials.

Like many scientific conceptualizations, the notion of assessment literacy is in need of careful explanation through the assessment and monitoring processes of language testing materials development. Assessment literacy is the extent to which learners, teachers and other stakeholders have familiarity and awareness of the assessment and its application in educational or academic contexts (Malone, 2013; Inbar-Lourie, 2008). It also refers to the principles, skills and knowledge that are required for every stakeholder to be able to check the degree to which the learning objectives have been achieved (Fulcher, 2012). Accordingly, assessment literacy acts as one of the effective precursors of learning (Malone, 2013).

A very recent and widely recognized model that has been proposed for curriculum development with consideration of assessment as one of its core components is Nation and Macalister's (2010) model. In their model, Nation and Macalister (2010) commit to the issue of materials development by referring to 7 main factors. These factors are namely, attention to environment, consideration of needs, attention to the methods and principles of materials development, focus on the ultimate objectives, concentration on the required contents and their sequences of representation, attention to format and type of representation of the materials, focus on monitoring and assessment procedures and finally attention to the process of evaluation (Nation, & Macalister, 2010).



Given the monitoring and assessment guidelines provided in Nation and Macalister's (2010) framework of curriculum design, a series of considerations are required. Table 1 presents different components which play an influential role during the monitoring and assessment process:

**Table 1. Monitoring and Assessment Guidelines, Adapted from Nation and Macalister (2010), p. 109.**

	Learners	The learners should feel good about their progress
	Teachers	The teacher should be able to assess and correct the outcomes to the activities.
<b>ENVIRONMENT</b>		The teacher should have time to assess and correct the outcomes to the activities.
	Situation	The assessment should be economical.
<b>NEEDS</b>		The assessment may need to match with the external testing set by the education department, a government ministry or testing agency.
	Lacks	The course should show that the learners are increasing their knowledge of the language.
	Wants	The course should show the learners that they are learning to do what they want to do.
	Necessities	The course should show that the learners are getting better at tasks they will need to do after the course.
<b>PRINCIPLES</b>		As much as possible, the learners should be interested and excited about learning the language and they should come to value this learning.
		Learners should receive helpful feedback which will allow them to improve the quality of their language use.

The outer layers of Nation and Macalister's (2010) model are the environment where a course is held, the needs of the learners, and the teaching and learning principles.

Successive changes in the choice of textbooks are to a large extent affected by acts of monitoring. According to Oxford dictionary (2010), monitoring refers to the process of observing or checking the development or status of something during a time period. As a matter of fact, monitoring is being favored for successful judgment for the choice and grading of textbooks (Lee, 2013, Widodo, 2007). In this regard, the concept of assessment literacy which is an aspect being prone to be investigated in the monitoring and assessment process of materials development should not be mistakenly considered as parallel to the assessment and monitoring dimension of materials development. In other words, one can go through the process of assessment and monitoring the testing textbooks in terms of the extent to which they have dealt with the process of assessment for the purpose of promoting learners' and teachers' levels of assessment literacy and this is what the present study aims to do.

Webb (2002) argues that assessment literacy underpins conscious awareness of the means for determining the knowledge levels of the learners, taking in the results of assessment procedures, and successfully employing the obtained results. In this sense, assessment is one method for monitoring students' learning achievements (Webb, 2002). One of the main principles of assessment is navigating the learning performance. In other words, if a stakeholder decides to promote the learning outcomes of the learners, he or she should first of all revise his or her assessment performances (Norton, 2009). The other influential factors that can affect the quality of assessment are application of new technologies, use of authentic materials and integration of different skills (Norton, 2009).

A fruitful combination of different criteria proposed by Tomlinson (1998) act as a good starting point for monitoring dimension of materials choice for teaching and learning purposes. In Tomlinson's (1998) framework, a series of criteria are required to be explored to determine the adequacy of a teaching or learning material. Examples of such criteria include influence, enjoyment, confidence, relevance, availability, authenticity, and self-investment (Tomlinson, 1998). Moreover, assessment is required to determine why, what, who, where and when questions regarding acts of monitoring (Cooper, 1997). More importantly, needs variations across various learning and teaching contexts increase the necessity of materials and textbooks are not exceptions of these materials (Harsono, 2007).

Success in monitoring acts strongly correlates with teachers' potential contributions. As Widodo (2007) argues ones of the causal roles of teachers in materials development procedures is monitoring whereby the teacher fixes the students



performance problems in different learning activities. Furthermore, Davison (1976) has considered teachers as the first and main supporters of every learning attempt. Consequently, teachers' decisions act as a precise estimate of the appropriateness of the textbooks (Lee, 2013). Provided with such accounts concerning the roles of teachers in language learning, Anstey and Bull (2004) suggest teachers to go through a continuous process of monitoring and learning assessment. Thereby, achieving knowledge about assessment helps teachers gain awareness of students' needs (Watson, 2010). What makes matters is the fact that future teachers are the current learners and attention should be directed towards the assessment needs of these learners.

From the perspective of Black, Harrison, Lee, Marshall, and Wiliam (2004), assessment and monitoring for learning are distinguished from other types of assessment such as assessment for accountability or ensuring competence. Black et al (2004) consider assessment for learning as a type of assessment whose design and practice processes revolve around promoting students' learning. Accordingly, assessment for learning yields feedback and opportunities for teachers to monitor their teaching performances (Black et al, 2004). Knowledge about such kinds of assessment directs scholars towards identification and satisfaction of learning needs (Black, et al, 2004).

Anstey and Bull (2004) highlight the link between assessment and monitoring and maintain that monitoring is the bridge between assessment and evaluation or a link between teaching and learning. In their model of monitoring, Anstey and Bull (2004) consider monitoring as the core, and assessment, practice and translation as the four components of monitoring and planning. Watson (2010) maintains that monitoring acts as the key determiner of teachers' decision making attempts. As suggested by Ozturk (2013), monitoring and assessment are the incorporated variables of the models within the scopes of goals of materials development.

As Norton (2009) maintains assessment should be embedded in curriculum development. One of the main deficits of the existing programs of materials development is initial focus on content coverage rather than learning promotion (Norton, 2009). Norton (2009) believes that incorporating assessment in the curriculum design should be prior to content consideration. In other words, the initial stage of materials development should be focus on assessment for the sake of learning development. One fundamental reason for the significance of going after the concept of assessment is that to date, few number of research studies have been conducted to achieve an appropriate state of knowledge of assessment in the learning contexts (Norton, 2009). As such, it is of special importance for the Iranian context which is subject to recent assessment studies.

## **1. METHOD**

### **1.1. PARTICIPANTS**

Participants of the present study were 32 learners studying English literature and English translation at BA level whose testing courses revolved around the book "testing language skills: from theory to practice". The age range of subjects was between 18 to 25 years. These participants belonged to a group comprised of English literature and English translation fields each with 16 students who were mainly from University of Isfahan and Shiraz University. In addition, 5 teachers having experience of teaching BA testing courses were involved for the textbook selection stage. Their average teaching experiences was 5.2 years.

### **1.2. INSTRUMENTS**

The main instrument of the present study was Farhady et al's (2007) book entitled "testing language skills: from theory to practice" which is the most applicable and well-known book for increasing students' assessment literacy needs during BA testing courses in Iranian Universities. The second instrument was a carefully designed interview whose questions were based on the discretion of university professors teaching testing courses.

### **1.3. PROCEDURES**

The underlying model for the current study was Nation and Macalister's (2010) framework of curriculum design. Yet, considering the issues of manageability and economy, the focus was narrowed down and centered on testing courses and widely applied textbooks at BA university level in the area of English as a foreign language learning to signal the need for testing textbooks which promote university students' assessment literacy. The main methodology which was applied for so doing was content analysis of the book with focus on the core concepts in accordance with Nation and Macalister's (2010) framework for monitoring the course curriculum.

Consulting with 5 experts professionalized in the area of language testing, the book "testing language skills: from theory to practice" by Farhady, et al (2007) was approved to be assessed and monitored in terms of the degree to which it could help students increase their assessment literacy. The analysis procedure was initiated by conducting interviews with BA university students who had passed their testing courses with the aid of Farhady et al's (2007) book on language testing. During the interview process, learners were free to talk about their specific needs to be met in testing courses and the extent to which their needs had been satisfied. After identifying the main assessment literacy needs of the learners, Farhady et al's (2007) book was assessed in terms of each of the identified assessment needs of the learners.

## **2. FINDINGS AND DISCUSSION**

In response to the question asking about their expectations of BA testing courses, the participants provided a number of answers. Learners' knowledge wants showed a high degree of overlap and facilitated coding and categorization of the data. The related numbers and percentages of the learners addressing their expectations and wants were identified and reported in Table 2:

**Table 2. Learners' Assessment Knowledge and Practice Requirements**

Required knowledge (content) and practice	Frequency	Percentage
Knowledge about the quality of tests	24	75
knowledge about assessment and testing strategies	24	75
knowledge about testing different skills	31	96.9
knowledge about the ethical issues in assessment process	29	90.6
understanding about different types of assessment	32	100
Familiarity with different sources for increasing assessment literacy	28	87.5
Understanding the standards of assessment	32	100

Without recourse to the practical issues in the assessment process, above 75 percent of the subjects had dwelled on the knowledge and theoretical issues in their expressed opinions. Yet, going further through the learners' answers through semi-structured interviews, the researcher came up with the conclusion that the subjects were merging their theoretical and practical needs. In other words, by for example expressing reliability and validity as one of their requirements, learners had both theoretical and practical aspects of these two concepts in their minds. Within the following step of the study, the book "testing language skills: from theory to practice" was assessed in terms of each of the identified components from the perspective of assessment literacy. In the first place, analysis contained the same components that had been expressed by the learners.

## **2.1. Analysis Based on the Learners' Expectations**

### **2.1.1. Knowledge about Test Quality**

The authors have highlighted the significance of quality in language test development in chapter 6. Two of the major factors affecting the quality of a newly developed test are reliability and validity. These concepts together with different methods of their identification have been addressed in this chapter.

### **2.1.2. Knowledge about Assessment and Testing Strategies**

The main content of the book has been navigated by testing as the authors' primary concerns. Clearly, there is much further to go with regard to the assessment process. Nevertheless, line by line reading and analysis of the provided content lends credence to the idea that in view of the authors, testing and assessment are similar and incorporated into each other.

### **2.1.3. Knowledge about Testing Different Skills**

Chapters 8, 9, 10, 11, 12, 13, and 14 focus on teaching the learners about different types of items that may be appropriate for testing different skills. Chapter 8 maneuvers on testing vocabulary by introducing the issues and guidelines that should be considered in developing these tests. Chapter 9 revolves around testing grammatical categories and structures. Different activities and examples have been presented, as well, so that the learners can attain enough knowledge about grammatical assessment. Chapters 10, 11, 12, 13 and 14 respectively prepare learners for testing pronunciation, listening comprehension, oral production, reading comprehension, and writing ability.

### **2.1.4. Knowledge about the Ethical Issues in Assessment Process**

Unlike the claims of the authors concerning a wide coverage of the major issues in language testing, the ethical considerations in language testing which is an important issue has not been addressed at all.

### **2.1.5. Understanding about Different Types of Assessment**

By virtue of different functions or purposes of the tests, different types of tests have been introduced in the second chapter of the book. After illustrating a figure of the prognostic and evaluation of attainment as the functions of tests, different tests under each functional category have been introduced one by one. Employment of this top-down process promotes students' knowledge about first the general functional categories of tests and then their detailed specifications.

Chapter 16 also brings the readers' attention back to the previously introduced functional tests but this time in a more detailed way. There is no doubt that far more explanations were required for this very concept especially because this issue is to a large extent tuned with the contextual considerations and theoretical changes in the area of language testing. Though introducing different steps for developing a functional test, the authors have not provided any practical examples to provoke learners to enter the realm of functional test development and deal with its complexities. It should be noted however, that in the activity section the authors have compensated for this lack of enough consideration by asking the learners to analyze what they have read in the chapter and to apply them in practice.



### 2.1.6. Knowledge about Different Sources for Increasing Assessment Literacy

If not at all, the book has rarely considered suggestion of further references for those who are interested in extensive studies of the topics. This may be closely associated to what the authors have held in the early pages of the book concerning the broad coverage of all the issues that decreases the need for more language testing sources. What makes matters is not what the authors claim, but rather what they have really done and practiced during the development of the book. Taking some of aforementioned shortcomings of the book, one can come up with the conclusion that if the authors had suggested some references at the end of each chapter, many of the deficiencies of the content could be resolved.

### 2.1.7. Understanding the Standards of Assessment

Though not explicitly, the authors have referred to some of the psycholinguistic criteria for category membership of the tests and especially test items in the third chapter. As suggested by the authors, form is a true representative of the physical appearance of the tests. The main focus of chapter 3 is the structures and types of items as the key components of any test. It has been maintained that the nature and function of the test determines the types of its items. Three classifications of items have been presented each with their own related examples. It should be noted however, that no mention is made of the bases for these categorizations. In the later sections, the authors introduce psycholinguistic categorizations that center on psychological processes while answering the test items and also the linguistic principles that underlie the development of each item. Only in the middle of the section allocated to psycholinguistic categorization the following sentence highlights one of the standards and criteria for assessment:

“Therefore, items can have a variety of forms depending on the psychological processes and the linguistic mode of both their stems and the responses” (p. 38)

Understanding the whole section relies heavily on understanding this very statement. Hence, this sentence should have been noted earlier in the initial parts of this section in order for the readers not to be confused. The arguments of the present chapter have not much extended by the authors and it seems that they aimed at reducing the number of psychological processes that underlie the identification of different categories of tests.

## 2.2. Analysis Based on the Authors' Objectives

The authors of the book were skeptical of the satisfaction of Iranian language learners' goals associated with language testing area. Thus, they had suggested the present book as an alternative that reflected the specific demands of the Iranian EFL learners and accorded to the expectations of the learners and teachers. They believed that the present book was in line with the developments of language testing area and looked at language testing from different perspectives. The book was being favored by its authors because of its organization of ideas and issues, its vivid and intelligible illustration of the concepts and its broad coverage of the language testing area that diminished the need for learners' reference to further texts. The authors also pinpointed that the book aimed at fulfilling the needs of its target readers. Except for organization and language of presentation, the other issues including to the needs of the learners and coverage of the issues were addressed in the previous section.

### 2.2.1. Language of Presentation

About 95 percent of the learners believed that the content of the book was presented in a lucid manner. The researcher went through the readability analysis of the first paragraphs of chapter of the book to determine the extent to which learners' impressions could accord to the findings of the readability analysis. The results of such analysis have been portrayed in Table 3.

**Table 3. Readability Analysis of the First Paragraphs of the Book Chapters**

	Flesch-Kincaid Grade Level	Level	Word count	Interpretation
Chapter 1	12.6	College	115	Difficult to read
Chapter 2	10	Tenth grade	116	Fairly difficult to read
Chapter 3	9.1	Ninth grade	127	Fairly difficult to read
Chapter 4	9.9	Tenth grade	104	Fairly difficult to read
Chapter 5	8.6	Ninth grade	103	Average/standard
Chapter 6	11	Eleventh grade	120	Difficult to read
Chapter 7	12.8	college	113	Difficult to read
Chapter 8	9.3	Ninth grade	97	Fairly difficult to read
Chapter 9	10.1	Tenth grade	115	Fairly difficult to read
Chapter 10	10	Tenth grade	122	Fairly difficult to read
Chapter 11	9.9	Tenth grade	136	Fairly difficult to read



Chapter 12	11.4	Eleventh grade	125	Difficult to read
Chapter 13	9.4	Ninth grade	119	Fairly difficult to read
Chapter 14	11.5	Twelfth grade	125	Difficult to read
Chapter 15	10.7	Eleventh grade	114	Difficult to read
Chapter 16	9.5	Tenth grade	106	Standard/ average

According to Table 3, none of the analyzed paragraphs are beyond the comprehension level of the learners. In other words, considering ages 17 and 18 as the college range ages, the subjects of the present study could all understand the presented language and in this sense, their judgments about the intelligibility of the language holds true.

### 2.2.2. Organization of the Book

No textbook can be immune to the criticisms of those who use it especially those who rely on it as a good source of knowledge. The organization of the present book is not free from deficiencies. The location of chapter 7 is not much favorable, due to the influential role of background knowledge on learners' understanding of the newly presented information.

The preliminary sentences of chapter 7 direct the readers' attention to the idea that linguistic development necessitates testing development. Four testing trends have been identified on the bases of which are different approaches of testing. Different theories of language testing have been introduced to support the notion that based on different needs and purposes or functions, different types of testing approaches should yield applications. Furthermore, it refers to different ways of looking at the concept of testing. No absolute categories have been determined, but rather tests are placed within a continuum. In other words, one cannot claim that there is an absolutely discrete point or integrative approach. In the integrative approach, context is taken into consideration, but this context does not have to do with the learners but rather to the form of representation of the items (i.e., contextualized items). In the functional approach, the context that is considered is mainly related to the learners including their needs, what they have learned, their levels of competence and so forth. According to the authors depending on the nature of the test, a specific approach should be applied. For example, in testing vocabulary, discrete point items are more favored.

One of the demerits of chapter 7 is that the authors have only introduced some of the important terms such as pragmatic tests, without any further explanations. Yet, in the activity section, a question compensates for the lack of enough provided information about pragmatic tests and the differences between pragmatic and functional tests. It seems that the authors have tried to simplify the expressed issues in a way that they could be understood not perfectly but to a possible extent. The focus of this chapter is on theories especially approaches that are based on these theories. No detailed mention is made of the theories, but focus revolves around approaches that are based on the four introduced trends and theories. The end of the chapter like other chapters is without any references to guide learners in increasing their knowledge about the issues discussed before. Bearing all the contents expressed in chapter 7, it seems reasonable to have chapter 7 just after the first chapter.

Taking the first chapter of the book into account, the reader may conclude that the first chapter takes the form of a preliminary to prepare its target readers for the issues that are being discussed in the following chapters. These early pages of the book revolve around the importance of tests in language learning. Addressing the differences between test, evaluation and measurement, the authors signal the need for innovative schedules for testing purposes. The authors refer to the concepts of testing, measurement and evaluation with no mention of assessment. Considering the provided definitions for testing, measurement and evaluation, the discussions provided in the coming chapters and the date of publication of the book, it can be implied that the boundaries between the concepts of assessment and testing have not been blurred. Additionally, the balance between performance and goals is introduced as an important aspect in the evaluation process to enable decision makers to judge about acceptability of one's performance.

In chapter 1 two sections have been allocated to the importance of tests: one under "language testing" title and the other by "why tests" title. Line by line reading of these two sections which are both on the merits of testing leads to the conclusion that it would be unreasonable to have two separate sections for the benefits of testing. The next parts present some brief information about different types of tests such as traditional, standardized and teacher-made ones. What makes matters here is that the authors lend credence to the idea that multiple choice items are placed within the category of traditional objective tests, yet discussion of multiple choice items is in a separate section, not a subsection of traditional testing. Furthermore, the philosophy behind writing about testing communication as the last section of the chapter is unspecified.

In the final section of each chapter some activities are provided. Chapter 1 includes five activities that if not all, most of which need studying the coming chapters for correct answers. Yet, incoherent package of information diminishes its usefulness. The suggestion for more applicability of the first chapter is that the authors write a preview of the coming chapters to provide enough backgrounds for readers and make them ready for the next concepts.

Unlike chapters 8 to 14 whose focus are on different language skills, chapter 15 correctly signals the issues related to cloze and dictation tests. In this chapter, Fog readability formula is also introduced though not in details, as a benchmark for determining the level of difficulty of the text. Since the issues mentioned in chapter 15 have been previously addressed more or less in the other chapters, the readers of the book might expect this chapter to be placed before chapter 8.



Another important aspect of the book organization is its activity sections. The present book does not follow a fixed method for developing the activity sections. Accordingly, in some chapters, activity sections serve as an assessment aid and in the other chapters; they guide the learners towards new contents. For example, in the first chapter, learners are directed towards new contents that are going to be addressed in the coming chapters. Yet, unlike the first chapter, the activity section of the second chapter is analytic and challenging and makes students think and discuss further on the issue. Other examples include the activity section of chapter 6 which is mainly concerned with the learners' attained knowledge from the chapter with no challenging issues to be discussed in the classroom. Additionally, the activity section of chapter 3 extends the discussions pinpointed throughout the chapter. This section is much thought provoking in terms of the extent to which it directs learners towards applying their understandings from the current chapter in practice.

### 2.2.3. Strengths and Weaknesses of the Book

Without further maneuvering on the already analyzed chapters, the remaining chapters are described in terms of their strengths and weaknesses as follows:

Without recourse to a brief introduction of inferential category of statistics, chapter 4 is almost exclusively about descriptive statistics. In the same vein, a seemingly perfect view of the descriptive statistics has been outlined. A principal limitation of this chapter is a limited reference to inferential statistics without offering any explanations for it. The strong tendency of the authors to focus on descriptive statistics has made them ignore the other category of statistical analysis or so-called inferential statistics. Most of the required information about descriptive statistics including measures of central tendency, and measures of variability has been satisfactorily addressed in this chapter. Further, different types of data tabulation and graphical illustrations of the data have been explained with examples. Finally, correlation is introduced as a required technique in the language testing area. A limitation which is again repeated in this chapter is presence of tables without caption which may confuse learners. The activity section of this chapter brings some questions based on the presented content of the chapter in order to help learners ensure that they do not have trouble with the issues discussed in this chapter.

Chapter 5 provides a detailed discussion of the process of test construction. The content of this chapter depends to a large extent on the previous chapters on the forms and functions of language tests. It is because the first step in test construction is identification of the forms and functions of the test. Then, a clear illustration of the other stages of test development is made. This helps readers achieve a state of knowledge about how to construct a test through understanding the main issues regarding the development of items and the related issues such as item facility and item discrimination. The activity section of the fifth chapter assesses learners' understanding of the previously discussed issues.

Furthermore, lack of captions or some inconsistencies for presenting the tables is a shortcoming of the book. The tables should have had captions in order to guide readers' attention to the specific issues they were tapping. Example of tables without captions is Table 3.1 on page 39. An appropriate caption for this table could be "psycholinguistic classification". There are also some other tables in chapter 4 that are presented without captions.

## 3. CONCLUSION

As with other learning areas, language testing is faced with many innovative theories. This has made the most applicable testing textbooks in need of evaluation. The process of evaluation is strongly correlated with exploring the conditions under which the textbooks are applied. Thus, evaluation should then be represented in such a way that the ideas of learners, teachers, and materials developers (in this study, author (s)) can be monitored side by side. An analysis of the extent to which objectives and expectations of these three main categories of stakeholders match, was the ultimate goal of the present evaluation. In so doing interviews were conducted to identify language testing and assessment needs of the learners. In addition, teachers were interviewed with regard to their preferences in selecting their testing course books and their purposes. Given the authors claims provided in the early pages of the book and the two categories of information obtained from interviews, the present analysis was navigated.

Evaluation of the book was instantiated with a fruitful combination of the book authors, learners, and teachers' opinions. For the sake of economy and precision of the findings, only the monitoring and assessment layer of Nation and Macalister's (2010) framework for curriculum development was taken into account. In this framework, monitoring is considered within the inner circle of course evaluation cycle. Thus, any attempts to evaluate materials - including textbooks- should incorporate the three outside layers (i.e., environment, needs and principles) that are comprised of monitoring and assessment, format and presentation and content and sequencing layers of this framework, as well.

Going through the book and seeking for learners' required components of assessment literacy, the researchers came up with the following conclusions:

The book addresses most of required components of assessment literacy for learners studying English literature and English translation at BA level in the Iranian context. Yet, there are some important issues such as ethical considerations in assessment, different sources of assessment and to some extent standards of assessment that have not been addressed in the book.

Except for the aforementioned shortcomings such as no coverage of all the issues and learner needs, the book has met the objectives mentioned by its authors including an acceptable and easy way of presenting the issues.

All in all, the book is very old and its references are not much up to date. Accordingly it has not provided information in line with the changes in the testing theories. Thus, due to the ever changing needs of the learners and society, teachers should equip their classrooms and students with materials that are up to date and in line with the current changes. This





does not mean that the use of classic texts is not recommended, but rather the classical texts should be applied together with the recently authored ones.

## REFERENCES

1. Anstey, M., & Bull, G. (2004). *The literacy labyrinth*. NSW: Pearson Prentice Hall.
2. Atkinson, D. (2008). *Investigating expertise in textbook writing: Insights from a case study of an experienced materials writer at work*. Paper presented at the Lancaster postgraduate conference in linguistics and language teaching. Retrieved June 3, 2014 from <http://www.ling.lancs.ac.uk/pgconference/v02/01-Atkinson.pdf>.
3. Black, P., Harrison, C., Lee, C., Marshall, B., & Wiliam, D. (2004). *Phi Delta Kappan*, 86(1), 8-21.
4. Cooper, J.D. (1997). *Literacy: Helping children construct meaning*. Boston: Houghton Mifflin.
5. Davison, W.F. (1976). Factors in evaluating and selecting texts for the foreign-language classroom. *ELT Journal*, 30(4), 310-314.
6. Farhady, H., Jafarpour, A. J., & Birjandi, P. (2007). *Testing language skills: From theory to practice*. Tehran: SAMT.
7. Fulcher, G. (2012). Assessment literacy for the language classroom. , *Language Assessment Quarterly*, 9(2), 113- 132: doi: 10.1080/15434303.2011.642041
8. Harsono, Y. M. (2007). Developing learning materials for specific purposes. *TEFLIN Journal*, 18(2), 169-179.
9. Inbar- Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing*, 25, 385-402.
10. 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.
11. Mahmood, K. (2010). Textbook evaluation in Pakistan: Issue of conformity to the national curriculum guidelines. *Bulletin of Education and Research*, 32(1), 15-36.
12. Malone, M. E. (2013). The essentials of assessment literacy: Contrasts between testers and users. *Language Testing*, 30(3), 329-344. doi: 10.1177/0265532213480129
13. Nation, I. S. P., & Macalister, J. (2010). *Language curriculum design*. New York: Routledge.
14. Norton, L. (2009). Assessing student learning. In H. Fry, S. Ketteridge, S. Marshall(Eds). *A handbook for teaching and learning in higher education: Enhancing academic practice* (pp. 32-150). New York: Routledge.
15. Nunez Pardo, A., & Tellez Tellez, M. F. (2009). ELT materials: The key to fostering effective teaching and learning settings. *PROFILE*, 11(2), 171-186.
16. Ozturk, G. (2013). A negotiated syllabus: Potential advantages and drawbacks in English preparatory programs at universities. *International Journal on New Trends in Education and Their Implications*, 4(2), 35-40.
17. Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory Into Practice*, 48, 4-11. doi: 10.1080/00405840802577536
18. Scarino, A. (2013). Language assessment literacy as self-awareness: Understanding the role of interpretation in assessment and in teacher learning. *Language Testing*, 30(3), 309-327. doi: : 10.1177/0265532213480128
19. Stevenson, A. (Ed), (2010). *Oxford dictionary of English*. Oxford: Oxford University Press.
20. Tomlinson, B. (1998). In B., Tomlinson (Ed.), *Materials development in language teaching* (pp.1-35). Cambridge: Cambridge University Press.
21. Watson, R. (2010). *Linking assessment and pedagogy: Pre-primary teachers' literacy practices in Catholic schools, Western Australia*. Doctoral Dissertation, Murdoch University.
22. Webb, N. L. (2002). *Assessment literacy in a standards-based urban education setting*. Paper presented to the American Educational Research Association Annual Meeting in New Orleans, Louisiana.
23. Widodo, H. P. (2007). Textbook analysis on college academic writing. *TEFLIN Journal*, 18(2), 109-122.

## Appendix: Learners' Interview Questions

1. Which book did you study for your language testing course?
2. What do you think a university student should learn about assessment?
3. What are the main components of assessment that are required for the Iranian learners?
4. To what extent do the available books prepare you for assessment literacy? What components should be included in the available testing books?