

PRINCIPLES FOR DEVISING A READING COMPREHENSION TEST: A LIBRARY BASED REVIEW

***Muhammad Ahmad**

****Dr. Aleem Shakir**

*****Muhammad Aqeel**

******Ali Raza Siddique**

Abstract

This study aimed to review principles for devising reading comprehension tests. It focused on format, duration, marking and tasks for the development of reading comprehension tests. For this purpose, the material was downloaded from online sources including web search engines, databases and journals which comprised of a number of books, research articles and power point presentations that were read and relevant information was extracted in a separate MS word file. In this process, Mendeley Desktop was also utilized to locate most relevant information from different articles and books. The data were searched with help of keywords. Other than keywords, the information was gathered by typing questions and phrases in the search bars of the aforesaid databases and web search engines. To have an overview of the current practices regarding reading comprehension tests, IELTS, TOEFL, TOEIC and GRE tests were also reviewed and through this process, different points of similarities and dissimilarities were observed. Through this process, it came to know that reading comprehension tests were very important since, they measured varied aspects of the reading skill. It also unveiled certain facts, strengths and weaknesses of the reading tests. Plus, the study explored certain principles and suggested that reading tests should be devised in the light of those principles.

Key words: reading comprehension; reading comprehension tests; reading comprehension test construction; reading test construction principles

* SSE English, Government High School, Hujra Shah Muqem, Okara, Pakistan

** Assistant Professor, Department of Applied Linguistics, Government College University, Faisalabad, Pakistan

*** Visiting Lecturer, Bahauddin Zakariya University, Multan, Pakistan

**** Department of Applied Linguistics, Government College University, Faisalabad, Pakistan

1. Introduction

Comprehension of reading and reading are two separate things. Reading skill develops in teacher centered classrooms and means decoding and translation of a text into spoken words or sounds. Whereas, comprehension of reading means understanding what is read. In simple words, comprehension of reading means the ability to read, understand and recall what is read (Rutzler, 2017). Comprehension of reading is very important (Yuill & Oakhill, 1991). Without comprehending, reading is an exasperating or useless activity of word calling. It is doubtless to say that a well-developed ability among the learners to understand what they read has a positive effect on their entire lives. The major goal of a reading test is to help learners develop the skills, knowledge, experiences (Texas Educational Agency, 2002).

When a student begins learning how to read, it is extremely important that regular reading progress monitoring tests are conducted so that it can be assessed whether the learner is reaching the age-appropriate milestones or not. This will ensure that the student is making steady progress in reading and is on track for independent reading (Tompkins, 2011). Thus, realizing the significance of reading comprehension tests, this study aims to extract principles for the construction of appropriate reading comprehension tests by reviewing prominent researches available online. This study, therefore, aims to seek the answers of following questions:

1. What is the purpose of reading comprehension tests?
2. What are the important facts about reading comprehension tests to be known to the teachers, learners and their parents?
3. What type of content does a typical reading comprehension test involve?
4. Which skills do the reading comprehension tests measure?
5. What are the different levels of reading comprehension?
6. What kinds of activities are involved in a typical reading comprehension test?
7. From which sources, should the texts of reading comprehension tests be selected?
8. What can be the different ways to test reading comprehension?
9. Why are most of the reading comprehension tests restricted to MCQs or fill in the blanks only?
10. What should be the speed for reading a paragraph in a reading comprehension test?

11. Can a cloze test be used to measure reading comprehension?
12. What should be the number of questions in a reading comprehension test?
13. What are the general principles for reading comprehension tests?
14. What are the limitations of a reading comprehension test?
15. How reading comprehension tasks should be designed?

2. Methodology

This is, indeed, a library based research and is qualitative as well as descriptive in nature. Therefore, the material for this essay has been downloaded from online sources i.e.: (1) Web search engines like Google Search, Google Scholar and iSEEK; (2) Databases like ERIC, JSTORE, SlideShare; and (3) Research Journals like Journal of Reading, Scientific Studies of Reading, Journal of Research in Reading, Reading Research Quarterly and Reading in a Foreign Language. The names of the said journals were obtained from Master Journals List (Clarivate Analytics) and SJR (Scimago). The reason for selecting journals from these two sources is that they are recognized being most credible and authentic sources of indexing academic research journals.

2.1 Data Search Procedure

The data for this study were obtained through above web search engines, databases and research journals. For this purpose, following procedure was adopted:

2.1.1 Keywords' Search

To search relevant information from databases, web search engines and journals, keyword technique was utilized. For this purpose, keywords covering the topic of the study were determined which included general and specific words or terms. Among general terms included reading, testing, evaluation, format, assessment, comprehension, testing, principles etc. Then, these terms were joined together with one another to make them specific to the topic of the essay like reading comprehension, reading comprehension test(s), testing reading comprehension, principles for devising a reading comprehension test, format of a reading comprehension test, assessing reading comprehension, assessment of reading comprehension etc. These terms and

keywords were typed in the search bars of different web search engines, journals and databases in two different ways. First of all, they were typed word by word in search bars and then they were retyped adding (+) like (reading + comprehension + tests + principles).

2.1.2 Research Question Search

Other than keywords, the information was also gathered by typing research questions in the search bars of the mentioned databases and search engines like ‘what are the principles for the construction of a reading comprehension test?’ or ‘what should be the format of a reading comprehension test?’ etc.

2.1.3 Phrase Search

After keyword and research question search, the information was sought through typing phrases in the search bars like ‘format of a reading comprehension test’, ‘principles for the construction of a reading comprehension test’ etc.

2.2 Data Processing

The said data, which comprised of a number of online books, research articles and ppt presentations, at first, underwent the process of abstract analysis which helped filter irrelevant data. The rest of data, which was relevant to the topic of this essay, was read and relevant information was extracted in a separate MS word file. In this process, Mendeley Desktop was also utilized to access most relevant information as well as references from different articles with the help of keywords.

2.3 Test Review

To have an overview of the current practices regarding reading comprehension tests, IELTS, TOEFL, TOEIC and GRE tests were also reviewed and through this process, different points of similarities and dissimilarities were observed. Common practices in these tests were discussed to seek principles for devising reading comprehension tests.

3. Observations and Discussion

3.1 Purpose of Reading Comprehension Tests

Past theorists as well as researchers have given great significance to the concepts of efficiency and speed in their thoughts about reading ability (Carver, 1992; Jackson & McClelland, 1979; Perfetti, 1985; Stanovich, 1986, 2009). For example, Perfetti (1985, p. 12) claims “the definition of reading ability comes from considering both speed and comprehension”. Therefore, reading comprehension tests intend to measure learners’ ability of reading and understanding academic as well as non-academic texts. In other words, the reading comprehension test is designed to measure how well a learner understands what he reads (Brown, 2004; Mohamad, 1999). Thus, it can be said that the basic purpose of reading comprehension tests is to test learners’ ability of understanding of what they read (Rutzler, 2017). The same is evident from the conventional practice of IELTS, TOEFL, TOEIC, and GRE. All of these tests check learners’ ability to comprehend what they read. Hence, a reading comprehension test should be designed to measure learners’ ability of understanding what they read.

3.2 Important Facts about Reading Comprehension Tests

Hirsch (2006) outlines a number of facts about reading tests which are unknown but are much important for the learners, teachers and parents to know for educational improvement. These facts include (1) interpretation and analysis of the information from textbooks or nonfiction books, graphs, reference materials, charts, audio/video presentations, electronic databases, diagrams, oral interviews etc. (2) comparison or synthesis of information from diverse sources, (3) distinguishing between irrelevant or relevant information, (4) relating new information to prior experience and knowledge, (5) understanding and use of the features of a text which make the information useful (such as level of diction, sequence, format) etc. This process, according to Hirsch (2006) helps the learners; (a) read the texts and thereby determine the main ideas, (b) identify personal preferences concerned with reading of nonfiction or fiction, (c) read and organize information for different purposes like report making, interviews, test taking, and task performing and (d) recognition of difference between opinions and facts. Thus,

to make the reading comprehension test useful, the students, their parents as well as teachers should be made aware of these facts.

3.3 Content of a Reading Comprehension Test

Reading ability is not a fixed or static ability. It has different functions of numerous aspects, involving content, content of the task, interest, knowledge, motivation, nature, text organization and features of the setting in which reading occurs (Lipson & Wixson, 1986). Therefore, a typical reading test involves varied content such as recognition of words or word groups, association of sounds with corresponding graphical symbols, understanding of explicitly stated information (e.g. I wish Anne had come=Anne did not come=my wish), understanding of relations within sentences (i.e. structure, fronting and theme, negation, complex embedding), understanding conceptual meaning (i.e. quantity / amount, definiteness / indefiniteness, comparison / degree, cause / result / reason / purpose / condition / addition / contrast etc.), anticipation/prediction (e.g. what will come next), identification of main idea/salient features of the text, making generalizations/drawing conclusions, making inferences, reading critically, deduction of word meanings through, understanding word formation (i.e. affixation, compounding, derivation, roots) (Aydin, 2013; Heaton, 1975). But Brown (2004) listed almost the same things in a different way. He organized the functions under two heads and called them micro and macro skills of reading.

3.4 Skills Measured by a Reading Comprehension Test

Different reading tests like IELTS, TOEFL, TOEIC, and GRE measure more or less different skills. For example, IELTS measures the skills concerning gist making, drawing main idea, reading for details, skimming, understanding logical argument and recognition of opinion, attitude or purpose. Similarly, TOEFL measures the learners' ability to comprehend academic reading materials. Likewise, TOEIC measures how well test takers understand written English and GRE measures learners' language efficiency by checking speed and accuracy.

3.5 Levels of Reading Comprehension

3.5.1 Levels Introduced by Mohamad

Mohamad (1999) describes that reading comprehension takes place at three levels i.e. (1) literal comprehension, (2) interpretive or referential comprehension and (3) critical reading. Level 1 is appropriate for vocabulary testing. At this level, surface meanings can be tested. The learners can be asked to locate ideas and information overtly stated in the texts. Similarly, at the second level, the students go beyond deeper meanings. They are thought to have an ability to read and analyze critically while reading. At this level, the learners can be asked tricky questions such as: re-arranging the ideas or topics, explaining writer's purpose, making summaries and drawing conclusions. Finally, at the third level, information is evaluated critically.

3.5.2 Levels Introduced by Brown

Brown (2004) has introduced two levels of reading comprehension i.e. (1) micro and (2) macro. At micro level reading comprehension tests enable the learners discriminate between graphemic as well as orthographic patterns, retain language chunks of varied lengths in STM, process writing at an efficient rate, recognition of words and interpretation of their pattern and recognition of grammar patterns, rules and word classes (e.g. noun, pronoun etc.). Whereas, on macro level these tests measure recognition and interpretation of rhetorical devices, recognition of communicative functions of the text, inference of the context, formation of ideas or events, inferring links between events, deduce cause and effect and to distinguish the literal and implied meaning (Brown, 2004). In comparison of the functions reported by Aydin (2013), Brown (2004) and Heaton (1975), the facts reported by Brown (2004) seem more reliable. The reason for it is that they are not only latest but also provide a well-structured view of information regarding the said functions. IELTS, TOEFL, TOEIC and GRE are also measuring reading at both micro and macro levels. Therefore, reading comprehension tests should be constructed to measure micro as well as macro level skills.

A comparison between suggestions by Brown (2004) and Mohamad (1999) reveals that both have certain strengths and weaknesses. For example Mohamad (1999) offers clear guidelines or principles for the measurement at different levels but it does not suggest proper

format for the testing of reading ability. On the other hand, Brown (2004) not only categorizes learning in different types but also suggests formats of items to test reading comprehension. Ideal seems to be if one constructs and conducts the test in the light of the suggestions of both of them.

3.6 Activities Involved in Reading Comprehension Tests

3.6.1 Activities Suggested by Heaton (1975)

It is not a testing of skills by reading aloud. Instead, it involves measuring of skills through silent reading. It involves two kinds of activities i.e. (1) intensive (reading of short extracts of average difficulty) and (2) extensive (reading of a complete chapter, book, and article). Most of the reading tests including TOEFL, TOEIC, IELTS and GRE etc. focus on intensive reading. The reason for it is that it is more economical to have a large number of items based on short reading extracts.

3.6.2 Activities Proposed by Brown

Brown (2004) introduced following types of reading to be measured in a reading comprehension test; (1) Perceptive reading: such type of reading is carried out on large stretches, letters, words, punctuation (bottom up technique), (2) Selective reading: it involves the artifact of test formats to ensure recognition of discursive, lexical, and grammatical features, (3) Interactive reading : for, reading is a process of meaning negotiation; readers bring a set of schemata to the text to understand it. The intake of this process is the product of that interaction so short texts are used (4) Extensive reading: it involves the reading of texts extending more than one page like articles, essays, stories, books to be read out of class hours etc. It assesses global understanding of the texts. Brown (2004) suggests to measure reading ability through the evaluation of all these four types of reading. For it, he gives proper guidelines which will be discussed in upcoming paragraphs.

3.7 Sources for Text Selection

Along with usual reading extracts, newspaper articles, literary texts, instructions from manuals of appliance or machinery, public notices, advertisements, maps, directory extracts, timetables etc. can be used as the content texts to test the reading skill. It will provide realistic

means of assessment as well as motivate the learners to know how target language is used in lifelike situations (Heaton, 1975). But Brown (2004) comes forward to present these points in more detailed and explicit way i.e. he arranges the texts for reading ability assessment into different types and enlists them under an umbrella term called genre of reading which involves (1) Academic reading: it entails texts of general interest (e.g. magazine, newspaper articles), technical reports (e.g. lab reports) dictionaries, directions, essays, opinion writing, papers, thesis, textbooks, (2) Job related reading: it includes applications, emails, financial documents, forms, labels, manuals, memos, messages, questionnaires, reports, schedules, signs and (3) Personal readings: such as emails, greeting cards, invitations, letters, lists, magazines, newspapers, notes, novels, jokes, short stories, drama, medical reports. Such lists, in the opinion of Brown (2004), enable the “readers to apply certain schemata that will assist them in extracting appropriate meaning” (p. 187). Popular international level tests such as IELTS, TOEFL, TOEIC, GRE etc. consult different sources differently. An IELTS test, selects narrative, descriptive, discursive or argumentative texts from books, magazines, journals, newspapers written particularly for a non-specialist audience. IELTS selects some non-verbal materials like graphs, diagrams, illustrations etc. Similarly, GRE selects material from textbooks, magazines, scholarly journals, books, biographies, books of literature from the fields of social science, arts, humanities, politics etc. In the same way, a TOEFL reading comprehension test constructor selects texts from a variety of academic topics, like history, art, science and psychology.

3.8 Ways to Test Reading Skill

There can be several ways of testing reading comprehension such as MCQs, open ended questions, short questions, cloze items (Cutting, 2017; Keenan, Betjemann & Oslen, 2008), true/false items, matching items, fill in the blanks ordering, summarizing etc. (Brown, 2004; Clapham, 2014; Heaton, 1975). IELTS measures reading skill with the help of MCQs, matching items, completion items, diagram labeling, summarizing, short questions etc. TOEFL uses MCQs, insertion of a sentence, completion of a summary and completion of a table in its reading comprehension test. It means that there is no fixed way to test reading comprehension. Therefore, a reading test constructor can adopt any of the given ways according to his requirement.

3.9 Restriction to MCQs and Fill in the Blanks

Pearson and Hamm (2005) summarized previous researches and reported that reading tests implicated many components which seemed to depend on formats supporting understanding. Thus, some tests rely on MCQs, or fill in the blank or cloze formats. The drive for high reliability mostly results in certain restrictions regarding the type of material to read and respond. Therefore MCQs, fill in the blank or cloze procedure can be a better option for a reading comprehension test.

3.10 Determination of Reading Speed

Before constructing an L2/FL reading comprehension test, L1 reading skill must be ascertained. Reading speed should also be determined for the passages. Usually, reading speed for an average reader is 200-300 words/minute and 300-500 words/minute for the fast one. However, it is realistic to include 300 words for a reading comprehension test. In a scanning test, the questions are given before reading the text to direct the readers to find specific piece of information whereas in skimming the students are allowed to write main points while reading the extracts (Heaton, 1975).

3.11 Cloze Procedure

Cloze procedure can be utilized as a suitable tool for reading comprehension measurement (Greene, 2001; Keenan, Betjemann & Oslen, 2008). However, the cloze procedure items should not be confused with fill in the blanks. In the former, deletion is made subjectively whereas in the latter deletion is carried out systematically. If seventh word is deleted in the first sentence, then every seventh word should be deleted in all the sentences. Fifth, sixth and seventh word for deletion are ideally most favoured. For a paragraph of 280-300 words 40/7 and for 500 words paragraph 40/12 deletions should be made. If you are including rearrangement items, it is advisable to provide students with one or two answers. If they put first or two answers incorrectly, they will possibly attempt the rest of the questions incorrectly. For completion items, there should be one correct answer only (Heaton, 1975).

3.12 Number of Questions

So far as the number of questions is concerned, Monroe Community College, State University of New York recommends and includes 20 questions for its reading comprehension tests. Some of them refer to the reading of paragraphs of different lengths. Some paragraphs are 75 words long while the others comprise of more than 75 words. The questions comprise of sentence-relationship types, in which the learners have to choose how a sentence is related with the other. However, a comparative review of TOEFL, IELTS, TOEIC and GRE reveals that that the said tests rely on different number of questions i.e. IELTS (40), TOEFL (20-30) and TOEIC (100). Thus, a test constructor can include 20-100 questions in a reading comprehension test.

3.13 Test Duration

So far as the duration of a reading comprehension test is concerned, the theorists have not specified any minimum or maximum limits. For this purpose, the person, going to conduct a reading comprehension test, can adjust time limit according to his need. Same practice has been observed in international level language tests. All of them are conducting reading comprehension tests with different durations. Such as IELTS gives 60 minutes, TOEFL gives 60-80 minutes and TOEIC reserves 75 minutes for a reading comprehension test. Thus, a good comprehension test might be of 60-80 minute duration. However, the test constructor can adjust time according to his requirement.

3.14 Number of Paragraphs

For MCQs items, real life based paragraphs should be given. It should be kept in mind that only intensive reading skills are to be measured. Therefore, short paragraphs should be given. The length for an extract ranges from 50-100 words for primary level, 200-300 words for elementary level and 400-600 words for advanced level. However True/false items may encourage guessing, therefore, such items should be based on rubrics i.e. each correct item will be awarded two marks and one mark will be deducted for every incorrect answer (Heaton, 1975). Heaton or any other researcher does not suggest any specific number of paragraphs. For this purpose, guideline can be taken from the present tests at international level. IELTS uses 3,

TOEFL uses 3-4 and GRE uses 1-5 paragraphs. Thus, 3-5 paragraphs might be given in a reading comprehension test.

3.15 General Guidelines and Principles

Brown (2004) gives general principles for the construction of reading comprehension tests which can also be taken as guidelines or suggestions i.e. (1) reading is a process of negotiating meaning therefore, the texts for a reading test should be interactive, (2) learners (to be tested) should have mastered bottom-up strategies (i.e. to assess separate letters, phrases, words) at initial levels top-down strategies (i.e. comprehension abilities) at the elementary and advance levels, (3) the texts (to be selected for the test) should support the development of formal schemata (i.e. background information and cultural experience to interpret it effectively) and (4) for, there is no such technology available as to measure what there actually is in the brain (it is unobservable and it has no product), reading comprehension assessment should be carried out through inference and (5) such texts be given as to enable learners to predict meaning.

3.16 Limitations of Reading Comprehension Tests

A number of researchers like Hirsch (2006, 2016), Kamhi and Catts (2017), Sternberg (1991) and Wattenberg (2016) have found different faults with reading comprehension tests. Sternberg argues that that the reading tests measure in a very narrow range of reading situations and tasks. This view suggests that the said test may be a measure of the limited set of skills in present assessment practices (1991). Reading test, in fact, is a sampling device. It does not test the complete range of knowledge domains. It uses some typical samples from a few domains, and students' performance on these samples is taken to estimate their reading comprehension (Hirsch, 2006).

Similarly, reading comprehension tests are basically the measures of topic specific knowledge, which implies that the learners, having more knowledge about different topics, will outperform the learners with less knowledge (Hirsch, 2016; Kamhi & Catts, 2017; Wattenberg, 2016). Similarly, Sternberg (1991) in his research levels some objections against reading comprehension tests saying that (1) reading tests measure massed learning, while most of the routine learning is distributed, (2) reading test recall is entirely intentional and not at all

incidental, (3) reading test recall is immediate, (4) reading in a typical reading test is usually for a single question type, typically multiple choice, (5) reading test passages tend to be emotionally neutral, (6) the reasoning in reading test passages is too good, (7) reading test passages are usually un-motivating and even boring, (8) reading tests measure evaluation but not the construction of arguments, (9) reading tests do not control for students' desire to read, (10) reading test situations minimize distractions and (11) reading tests are for a single purpose. These are serious faults with reading comprehension tests. However, the measurement of reading ability is also very important. Therefore, reading tests might be constructed in a way to minimize the effects of these faults to get good results.

3.17 Design of Reading Comprehension Tasks

So far as the designing of reading assessment tasks is concerned, this study as mentioned earlier relies partially on the framework, guidelines or suggestions given by Heaton (1975) and partially on Brown's (2004). The reason for this reliance is that both of the said writers have given clear cut principles and formats.

3.17.1 Construction of Tasks for Perceptive Reading

Such type of reading is tested at rudimentary levels. The learners at this level are thought to be at an early stage of becoming literate. Some of the learners, at this level, have learned to read L1, but sometimes L2 works as an L1 for them. Therefore, age and cognition related issues are to be kept in view while constructing the test (Brown, 2004; Cooper, Robinson, Slanksy & Kiger, 2014). For, it is an initial or basic level test; the items for this test should be given to recognize the symbols, capitalization or lower case letters and punctuation. These things can be analyzed in following ways like (1) Reading Aloud/Written Response: (it tests words/letters, sentences reading aloud separately) and (2) Multiple Choice: MCQs (with 4-5 answers) along with other items (same/different, encircling answers, matching or true/false) are suitable to measure reading at lower level.

3.17.2 Construction of Tasks for Selective Reading

It is carried out at elementary level with main focus on formal aspects of language (e.g. lexical and grammatical aspects). Possible tasks for this level include (1) MCQs (form focused:

focus on basic vocabulary forms), like, ‘The mouse is ----- the table’. (2) In MCQs (Context focused) darting from one context to another takes place like:

Manager: do you like to work by yourself?

Employee: Yes, I like to work-----

A. Independently

B. Impatiently

C. Definitely

Modified cloze is a better format for it.

Cloze formats can also be used at this level. But keep in mind that story’s context will not help testees respond to items easily. Therefore, give one set of related sentences for eight items. Such items give authenticity to assessment. Similarly, matching items can also be utilized at this level. Such items are suitable for vocabulary testing. Communicative quality to matching can be added by leaving blanks in the first list. It can offer an alternative to MCQs or fill in blanks. Finally, it can be used as puzzle solving by adding 10-20 items. Similarly, editing tasks can also be utilized to measure reading at this level. Such items are considered best and widely used for assessing grammatical or rhetorical errors. It is also in practice in TOEFL. Picture-cued tasks can be used at this level too. For it, testees can also be asked to read a sentence or paragraph and to choose a picture. The testees can be asked to read different definitions or sentences each describing a labeled part in the picture and then be asked to identify a particular label. Similarly, gap filling may also be used at this level because it involves both reading and writing. The testees are asked to read the phrase and complete it by providing missing information.

3.17.3 Construction of Tasks for Interactive Reading

It involves a combination of form and meaning focused objectives. It mainly emphasizes on meaning. Therefore, it applies top down technique. It selects little longer texts than paragraphs and complex format graphics. It can use following formats; (1) Cloze Tasks: these tasks measure the ability to fill in the gaps. Reader can be asked to close that gap with a

calculated guess. A cloze test is, basically, meant for native readers. Oller (1973, 1976, 1979) defends its use for second language learners as well. Minimum length of texts for cloze procedure should be two paragraphs. Normally, a passage may consist of 30-50 blanks. However, a passage with half dozen blanks may also be called cloze. There are two approaches for deletion (1) exact word deletion (2) appropriate word deletion (e.g. for gorgeous sunset, beautiful, amazing and spectacular can also be used). (3) Editing texts can also be used at this level. Passage should be 200-300 words long comprising of 32-56 items (Imao, 2001). In addition to these scanning and ordering task can also be used (Alderson, Clapham & Wall, 1995).

3.17.4 Construction of Tasks for Extensive Reading

It involves longer texts like articles, magazines, essays, books etc. Therefore, top down processing is more suitable for it. Major focus of testing at this level should be on meaning. It can be measured through short answer items, editing, scanning, ordering and information transfer.

4. Conclusion

Different studies give significance to the different skills like the measurement of speed, efficiency, comprehension, understanding etc. but the common practice in IELTS, TOEFL, TOEIC and GRE is the measurement of how well a learner understands what he reads. Therefore, a reading comprehension test should be devised on this principle. There are a number of important aspects or facts regarding reading comprehension tests which should be well known to the teachers, learners and their parents but none of them is aware of them. Resultantly, the tests are not being utilized effectively. Therefore, if someone wants to utilize reading comprehension tests more effectively, he will have to make teachers, learners and parents aware of these aspects. Sometimes, reading is taken as a fixed or static ability. But the fact is otherwise i.e. reading is a dynamic activity. For, it involves different functions of multiple aspects, like content, tasks, knowledge, interest, text organization, context, etc. therefore, if someone, wants to make a good comprehension test, he will have to involve varied contents to effectively measure reading as a dynamic activity but in accordance with the given principles.

IELTS test measures gist making, drawing main idea, reading for details, skimming, understanding logical argument, and recognition of opinion, attitude or purpose etc. Similarly, a TOEFL reading test measures the learners' ability to comprehend academic reading materials. Likewise, TOEIC measures how well test takers understand written English and GRE measures learners' language efficiency by checking speed and accuracy. Thus, it seems that different tests focus on different skills but one thing is common i.e. ability to understand while reading. Therefore, if someone wants to make an effective reading skill, he should focus on measuring learners' ability to comprehend while reading but the principles for devising an appropriate test should not be ignored. A comparison between suggestions given by Brown (2004) and Mohamad (1999) reveals that both have strengths and weaknesses. For example Mohamad (1999) offers clear guidelines or principles for measurement at different levels but it does not suggest proper format for the testing of reading ability. On the other hand, Brown (2004) not only categorizes learning in different types but also suggests formats of items to test reading. Ideal seems to be if one constructs and conducts the test in the light of the suggestions of both of them. Heaton (1975) has introduced two reading skills i.e. intensive and extensive but Brown (2004) has given four types of activities like perceptive, selective, interactive and intensive. Whatever, these skills are, IELTS, TOEFL, TOEIC and GRE are measuring intensive reading skills. Therefore, if someone intends to make a good reading comprehension test, he should focus on intensive skill only but in the light of given principles.

Different test agencies like British Council, ETS etc. and different researchers like Brown (2004) and Heaton (1975) have enlisted different contents for the selection of the text for a reading comprehension test including newspaper articles, literary texts, instructions from manuals of appliance or machinery, public notices, advertisements, maps, directory extracts, timetables etc. But point noteworthy here is that the content should not be the main focus of a test constructor. Instead, if someone wants to devise a good test for reading comprehension, he should focus on the text ensuring realistic and lifelike situations but he should not forget the principles of selecting the texts and framing them into the test. There are different ways of testing reading such as MCQs, open ended questions, short questions, cloze items, true-false items, matching items, fill in the blanks, summarizing etc. These are all available options for the construction of a reading test. But there is no compulsion to select all of these or any one of them. Instead, a test constructor is free to devise test including all or few of them. If he relies on

MCQs only and ignores the rest, his test will be effective but one thing should be kept in view i.e. the principles of that item. Different tests and researchers suggest different number of questions, different number of paragraphs and different duration for a reading test. There is no particular guideline regarding these things. Thus, a test constructor is free in these cases. But this freedom can affect the quality of his test. Therefore, if he wants to make his test effective, he should follow the proper principles. In addition, there can be many faults with a reading comprehension test. But they can be overcome, if the test is devised in the light of the principles.

References

1. Alderson, J. C., Clapham, C., & Wall, D. (1995). *Language test construction and evaluation*. Cambridge: Cambridge University Press.
2. Aydin, A. (2013). *Testing reading comprehension*. Retrieved from : <https://www.slideshare.net/aylin22/testing-reading-comprehension-21445232>
3. Brown, H. D. (2004). *Language assessment principles and classroom practices*. New York: Longman.
4. Carver, R. P. (1992). What do standardized tests of reading comprehension measure in terms of efficiency, accuracy, and rate? *Reading Research Quarterly*, 27(4), 347-359.
5. Clapham, C. (2014). *Assessing reading*. Retrieved from https://www.slideshare.net/mrssuarez/assessing-reading-40736845?next_slideshow=1
6. Cooper, J. D., Robinson, M. D., Slansky, J. A., & Kiger, N. D. (2014). *Literacy: Helping students construct meaning*. Australia, Brazil, Mexico, Singapore, UK and USA: Cengage Learning.
7. Cutting, L. E. (2017). *Do all reading comprehension tests assess equally?* Retrieved from <https://www.hmhco.com/blog/do-all-reading-comprehension-tests-assess-equally>
8. Greene, B. (2001). Testing reading comprehension of theoretical discourse with cloze. *Journal of Research in Reading*, 24(1), 82-98.
9. Heaton, J. B. (1975). *Writing English language tests: A practical guide for teachers of English as a second or foreign language*. London: Longman.
10. Hirsch Jr., E. D. (2006). What do reading comprehension tests measure. *Knowledge. American Educator*, 30(1). Retrieved from https://www.aft.org/pubsreports/american_educator/issues/spring06/tests.html

11. Hirsch Jr., E. D. (2016). *Why knowledge matters: Rescuing our children from failed educational theories*. Cambridge: Harvard Education Press.
12. Imao, Y. (2001). *Validating a new ESL placement test at SFSU*. (unpublished master's thesis). Department of English, San Francisco State University, San Francisco, America.
13. Jackson, M. D., & McClelland, J. L. (1979). Processing determinants of reading speed. *Journal of Experimental Psychology: General*, 108(2), 151-181.
14. Kamhi, A. G., & Catts, H. W. (2017). Epilogue: Reading comprehension is not a single ability—implications for assessment and instruction. *Language, Speech, and Hearing Services in Schools*, 48(2), 104-107.
15. Keenan, J. M., Betjemann, R. S., & Olson, R. K. (2008). Reading comprehension tests vary in the skills they assess: Differential dependence on decoding and oral comprehension. *Scientific Studies of Reading*, 12(3), 281-300.
16. Lipson, M. Y., & Wixson, K. K. (1986). Reading disability research: An interactionist perspective. *Review of Educational Research*, 56(1), 111-136.
17. Mohamad, A. (1999). What do we test when we test reading comprehension. *The Internet TESL Journal*, 5(12). Retrieved <http://iteslj.org/Techniques/Mohamad-TestingReading.html>
18. Oller, J. W. (1973). Cloze tests of second language proficiency and what they measure. *Language Learning*, 23(1), 105-118.
19. Oller, J. W. (1976). A program for language testing research. *Language Learning*, 26(1), 141- 166.
20. Oller, J. W. (1979). *Language tests at school: A pragmatic approach*. London: Longman.
21. Pearson, P. D., & Hamm, D. N. (2005). The assessment of reading comprehension: A review of practices-past, present, and future. In S. G. Paris & S. A. Stahl (Eds.), *Center for Improvement of Early Reading Achievement (CIERA). Children's Reading Comprehension and Assessment* (pp. 13–69). New Jersey: Lawrence Erlbaum Associates Publishers.
22. Perfetti, C. A. (1985). *Reading ability*. New York: Oxford University Press.
23. Rutzler, S. (2017). Importance of reading comprehension. *Math Genie*. Retrieved from <https://www.mathgenie.com/blog/importance-of-reading-comprehension>

24. Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-407.
25. Stanovich, K. E. (2009). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Journal of Education*, 189(1-2), 23-55.
26. Sternberg, R. J. (1991). Are we reading too much into reading comprehension tests? *Journal of Reading*, 34(7), 540-545.
27. Texas Educational Agency. (2002). *Comprehension instruction*, 4-8. Retrieved from http://www.netxy.net/pm_attach/67/TRI-Comprehension_Instr.pdf.
28. Tompkins, G. E. (2011). *Literacy in the early grades: A successful start for prek-4 readers* (3rd Edition), Boston: Pearson.
29. Yuill, N., & Oakhill, J. (1991). *Children's problems in text comprehension: An experimental investigation*. New York: Cambridge University Press.
30. Wattenberg, R. (2016). Inside the common core reading tests: Why the best prep is a knowledge-rich curriculum. *Knowledge Matters* (Issue Brief #7). Retrieved from <http://knowledgematterscampaign.org/wp-content/uploads/2016/09/Wattenberg.pdf>