3.0. Introduction:

In the previous chapter, a detail survey of the theoretical background and the related literature of the study was given. In this chapter, however, a detailed explanation of the methodology adopted in carrying out the research of the present study will be explained. Certainly, it is going to throw light on the sample of the study, tools designed for data collection, procedure, objectives and hypothesis on which the research is based.

3.1. Sample Design:

The subjects of this study were drawn from the population in a simple random sampling method or what so called ‘probability’ or ‘chance’ sampling, which enabled the researcher to make sure, as Kothari (1990:73) points out that “every item of the universe has an equal chance of inclusion in the study. It is, so to say, a lottery method in which individual units are picked up from the whole group not deliberately but by some mechanical process.” He further argues in favor of the reliability of this technique that this sort of sampling ensures the law of Statistical Regularity which demonstrates that if on a standard the sample selected is a random one, the sample will bear a resemblance to the universe in all features. This is why random sampling is regarded to be the best technique of sampling (ibid: 74).

23 subjects of the first year, henceforth group A, were chosen from a total number of 100 students in the rolls of the department. Likewise, 23 students of the second year, henceforth group B, were selected from a total number of 55 enrolled students of the department. For sampling purposes, each student in both first and second year was assigned a particular number. These numbers, certainly, were written on small slips and then each was folded in the same way and put in a small container and mixed together thoroughly. This was, unquestionably, done for each group separately. As mentioned above, 23 students from each year were drawn by a seven-year-old child in a lottery manner to eliminate bias and to ensure
that each member of the population would have the opportunity to be included in the study. Needless to mention, after every pull the whole slips were reshuffled again to make selection go on at the same manner of randomness.

It is worth mentioning, at this point, that in the present study the whole sample consisted of two homogeneous groups. Group A comprised 23 male students of approximately the same age, 20+ years, where as group B consisted of 23 male students 20+ years of age. What was really striking about the subjects of this study was the absence of the females. In fact, there were only 2 female students in the whole department, exactly in the first class, at that particular academic year in which this research was carried out. However, the number of girls in the college of education, Shabwa, was very limited in general. For further details on the girls’ education in Shabwa Governorate and in Yemen Republic as a whole see (1.2).

Six out of twelve of the teachers of the department participated in the study. The other remaining teachers were on study leave for both M.A. and Ph.D. courses at the time of the study. A questionnaire was designed and administered for the teachers. Out of the total number participated in the study there was only one female teacher. The teaching experience of the teachers ranged from two years to more than five years: 83.3% indicated that they had been teaching for more than five years, and probably for some of them experience exceeded fifteen years or more of teaching English, while 16.7% two to four years. Four of them had Bachelor in Education (B.Ed), one had Bachelor in Arts (B.A.); where as one had got a Master of Arts (M.A.).

3.2. Objectives of the Study:

The study has the following major objectives:

1) To find out what makes reading comprehension in English problematic for EFL learners in this particular department.

2) To locate the factors which obstruct or encourage the teaching of reading abilities and reading comprehension.
3) To observe to what extent the teachers' awareness of reading skills, and how they deal with teaching reading in terms of approaches and methods.
4) To check the relationship between habit of reading and readers' interests, on the one hand, and the performance of the learners in reading comprehension, on the other.
5) To look into the interaction of reading, as a receptive skill, with writing, as a productive skill.

3.3. **Hypothesis:**

The present study is based on the following hypotheses:

1) Reading an English text is difficult for EFL/ESL students who have not been acquainted with the appropriate reading skills and need to interpret or decode close or open words and to infer, and extract information from a given text.

2) Reading an English text is relatively not difficult for EFL/ESL learners who have been acquainted with the appropriate skills to decode and interpret a certain reading material and have the ability to infer and extract meaning from a certain text.

3) EFL/ESL students, who have been appropriately taught reading skills, can produce well arranged and good organized pieces of writing.

3.4. **Tools of the Data Collection:**

The tools of the data collected for the present study consisted of cloze tests, comprehension tests, composition tests, questionnaires, and observation of the actual teaching of reading. These will, in fact, be discussed in turn in the following paragraphs.

3.4.1. **Cloze Tests:**

Cloze procedure is defined by Harris et al (1983:53) as “any of several ways measuring a person's ability to restore omitted portions of an oral or written message from its remaining context.” The term itself was coined by
W. Taylor in 1953 to reflect the Gestalt principles of 'closure', the ability to complete an incomplete stimulus.

Types of cloze tests are noticeably classified by Alderson (1996:222). He makes a distinction between two types of cloze tests; the first one is pseudo-random cloze test in which every nth word is deleted. Construction procedure for this test insures that the test writer does not really know what is being tested. The more deletions are the larger chances are found for many aspects of language to be involved. The other type is rational cloze technique, or what is also known as fixed ratio cloze test in which the deletion is under the control of the test writer, i.e., tester decides which words to be omitted, and tries to leave at least five or six words between gaps (otherwise gaps are unduly difficult to restore).

As far as reading is concerned, a standardized procedure and set of scores are developed to differentiate frustration/instrumental reading levels and instructional/independent reading level. Connelly (1997) asserts that the principles of the cloze method function by the regular removal of every nth word in a test which is authentic in that it is not specifically created but taken from a source considered appropriately by the tester. Additionally, Anderson (1997:178) in discussing cloze tests as technique for measuring reading comprehension and readability argues that it is "a testing technique that has excited much interest research activity with English as a first language. Strangely, there are only few studies where the technique used with students for whom English is a second or foreign language." He also suggests that cloze procedure may also be utilized to measure comprehension abilities of non-native speakers learning English.

Wainman (1979:131) concludes that the cloze method of automatic deletion can attain results for second language learners which are as good as the other results achieved by the more popular methods, which usually favors the content words. Cloze tests are easier to set and they explore a wider variety of language categories than do those tests which concentrate on continuous texts or single sentences. Recently, there are some other advantages of cloze procedure identified by Weir (1990) and reported by Connelly (1997:142) as:
1) Easy to construct and score, yet they are valid indicators of language proficiency.

2) A relatively short text makes the test easy to handle yields a test with good internal consistency.

3) Both validity and uniformity are claimed for this type of test by its proponents.

Jonz (1987:414) argues that "...cloze, the practice of requiring examinees to restore deleted words to text, is well suited to the study of interaction between text and individual comprehension processes precisely because it is an integrative testing procedure. It requires examinees to apply knowledge at various levels of generality, ranging from the morphosyntax of phrases to broad concepts and systems of belief." Furthermore, the items on a typical every n<sup>th</sup> word or fixed ratio cloze process in Bachman' view (1985: 539) can be stored into four categories. The test-taker, however, must apply (1) within clause knowledge, (2) across clause, within sentences knowledge; (3) across sentence, within text knowledge; and (4) extra-textual knowledge.

More recently, Farhady et al (1996: 279) argue that various research projects were conducted to investigate the efficiency of cloze procedure as a test for non-native learners' language proficiency. Through the course of these projects, the concept of cloze became established with more precise and research supported characteristics.

Regarding the most appropriate value of "n", cloze tests with every 3<sup>rd</sup>, 4<sup>th</sup>, 5<sup>th</sup>, 6<sup>th</sup>, 7<sup>th</sup>, 8<sup>th</sup>, 9<sup>th</sup>, 10<sup>th</sup> word deletion were developed. However, test with every 5<sup>th</sup> or 7<sup>th</sup> word deletion have become more popular. Cloze test with every 7<sup>th</sup> word deleted is considered to be the standard cloze test (op. cit).

In discussing the total number of words deleted in a cloze test, Farhady et al (1996) argue that the appropriate deletions range between 20 – 100. But there are certain issues that remain contentious such as the length of the passage and the validity and reliability of such tests. They, however, discussed these matters further:

*Although a 25-item test may show an acceptable degree of validity, it does not, however, guarantee high validity. Therefore, the number*
of items should be determined at a level which would produce both high reliability and high validity. Experiments demonstrate that cloze tests with 25–30 items show as high a validity index as those with 40–50 items. Thus, common sense criterion for the number of deletions in a cloze test was set to be 25 to 30. It means that considering the 7th rate and 25 to 30 blanks, the length of the cloze passage would be somewhere between 175 to 210 words.

Farhady et al (1996:280)

Cloze tests, however, are as valid as any other type of tests. They are, certainly, very practical in several circumstances because they are uncomplicated to arrange and score; their validity as tests for reading is, some what contentious. Cripwell (1976) argues that “cloze tests are comparatively easy to construct. And also, in these tests student has to make up his mind” and it is argued that this is why this test unlike most other tests is creative rather than a failing test.

Regarding scoring cloze tests, Oller (1973) suggests that “the most frequently used with native speakers is the counting of ‘exact word replacement’ as correct”, yet, with non-native speakers there are some other methods among which is the ‘acceptable word scoring method’, which is, according to him, an acceptable and quite significant. Alderson (2000: 207) argues that in some scoring measures, credit may also be given for supplying a word that makes sense in the blank, even if it is not the word which was originally deleted.

In the present study, and for the above discussions, cloze tests were used to assess the learners’ readability and comprehensibility by selecting some passages from particular texts and they were developed for testing the cloze test items. These tests helped the researcher to a great extent to investigate the learners’ problems of reading in particular. Cloze tests, generally, can possibly be used to check the level of readability of a particular text, as said above. This factor is much related to the present study of assessing difficulties faced by foreign language learners reading an English prose text. If, however, learners show some difficulties in reading then it might be necessarily momentous to concentrate on those areas in which the learners’ experience reading problems and in the process assist them arrive at an understanding of the text.
3.4.2. Comprehension Tests:

Comprehension tests are of different types: multiple choice questions (MCQs), open-ended questions (OEQs), true/false questions (T/F Qs) or ordering task questions (OTQs).

Alderson (2000:211) argues that MCQs are a common device for testing students' text comprehension. They allow testers to control the range of potential answers to comprehension questions, and to some extent to control the students' thought processes when responding. He also cited Munbay (1968: xii) as saying:

Multiple-choice questioning can be used effectively to train a person’s ability to think....It is possible to set the distracters so close that the pupil has to examine each alternative very carefully before he can decide on the best answer....When a person answers a comprehension question incorrectly, the reason for his error may be intellectual or linguistic or a mixture of the two. Such errors can be analyzed and then classified so that questioning can be taken account of these areas of difficulty.

Alderson (2000:204)

Ingram (1974:322) identifies the following advantages of MCQs:

1- They are much quicker and more convenient to mark.

2- They are independent of the judgment of the maker, in contrast with open-ended items.

3- They provide more diagnostic information, because the candidate is forced to decide among a particular set of choices.

Open-ended questions (OEQs) are also used as a technique for measuring reading in which the test-taker writes his own answers in the form of a single word, a phrase or a sentence. Ingram (op. cit.) identifies the following three advantages of OEQs:

1) There is no possibility of anybody getting the right answer by chance.

2) Open-ended items test productive skills. Multiple-choice items test receptive skills only.

3) Open-ended items do not present the learner with errors, and so do not risk teaching wrong thing.

MCQs are best documented in the words of Mathew-Bresky as follows:
The usefulness of multiple choice language testing has been recognized for some time by teachers and examiners; collections of tests passages with accompanying questions have been published, courses are furnished with multiple-choice tests materials, and the provision of credible distracters drawn from the test passage or outside it has been raised to a fine art. But most development has been on multiple-choice as a device for testing comprehension, usually written but also aural, where as there is also wide scope for the use of multiple-choice to test strictly linguistic items at beginners and intermediate levels.

Mathew-Bresky (1978:50-51)

Furthermore, Nuttall (1996:186) in discussing questioning techniques in reading specifies the following advantages:

1) They are relatively easy to device.
2) They can be used for virtually any purpose.
3) They force the student to think things out for himself.

True/False questions (T/F Qs), moreover, present a statement; the students have to decide if it is true or false according to the text. Nuttall (op cit) argues that of the question types, T/FQ is in some respects the most limited, but it can be useful and is not necessarily trivial. T/FQs can promote discussion like MCQs; more choices make discussion more interesting, but T/FQs are preferable if there is one likely misinterpretation of the text. InT/FQs, according to Alderson (2000:222) test-takers are given a statement which is related to a target text and have to point out whether this is true or false, or whether the text agrees or disagrees with the statement.

Finally, another type of reading questions is ordering task question (OTQ) students are given scrambled set of words, sentences, paragraphs or texts, and have to put them in their correct order (See Alderson op. cit.).

Whatever technique is followed to measure reading, the main focus is to check the learners’ comprehension of a certain prose text. It is to measure their ability in certain sub-skills such as ‘inferring’ or what is commonly known as ‘reading between the lines’. The passages sampled for this study were selected from different texts to measure students’ ability in inference, their reading for detail, and their ability for handling vocabulary. Different techniques were used in achieving these goals such as MCQs, in which some questions in all the three passages for both groups A and B were
designed. OEQs were also used especially for the measuring students, reading for detail. Further, T/F and OTQs were used as techniques to elicit information about the learners’ reading comprehension in the present study.

In sampling the passages for testing the learners’ reading ability in this study a careful selection was followed. The three texts of which these six passages (total number of the passages), however, were once upon time used for reading classes in the department of English, Aden University. These passages were carefully chosen to suit, as far as possible, the level and the background knowledge of the examinees, i.e., to tackle topics which seem to be common to the learners and within the scope of their level of learning and background knowledge.

3.4.3. Composition Tests:

Recently, many scholars have argued in support of a close relationship between reading and writing: as skills, as cognitive processes, as ways of learning. Grabe (1991:394) argues that for sometime, L1 researchers point out the high correlation between good writers and good readers and view reading and writing as mutually reinforcing interactive processes. Stotsky (1983) cited in Grabe (1991:394) notes that better writers are better readers, better writers read more, and better readers write more syntactically mature prose, and reading experiences improve writing more than grammar instruction or further writing exercises.

The subjects of both group A and B were assigned two free composition tests, one for each. Group A subjects, for instance, were asked to write on the topic: “My work” whereas group B students were asked to compose an article on; “The village/ town/ city in which I live in”. The rationale behind the designing of such tests in composition is to check the students’ ability of writing. If they are really good at reading they are expected to perform well in writing. The composition tests are to examine their language proficiency and problems of language learning as well.

3.4.4. Observations:

Observation is the procedure in which one or more persons observe what is taking place in a number of real-life situations, and they categorize and record relevant activities according to some previously designed schedule
plans. In Best and Kahn's view (1993:222) as a data collection tool, direct observation may also make an important contribution to descriptive research. Certain types of information can best be attained all the way through direct assessment by the researcher. However, they further argue that when the observation concerns aspects of material objects or specimen, the process is relatively simple, and may consist of categorizing, assessing, or counting. But when the process engages the study of human subjects in action, it is much more multifaceted.

In addition to that, kothri describes the observation technique as:

*The observation method is the most commonly used method specially in studies relating to behavioral sciences....Observation becomes a scientific tool and the method of data collection for the researcher, when it serves a formulated research purpose, is systematically planned and recorded and is subjected to checks and controls on validity and reliability.*

*Kothri (1990:118)*

This method of collecting data is being criticized by some experts because it is expensive method; information provided by observation is limited, and sometimes unforeseen factors may intervene with the observational task. Nevertheless, if properly used, observation method according to koul (1997:175) has the following advantages:

- It provides a direct procedure for studying various aspects of human behaviour, which may be the only effective way to gather data in particular situation.
- It enables the observer to code and recode behaviour at the time of its occurrence.

In addition to these two advantages, Best and Kahn (1993:229) classified a number of characteristics of good observation of which the following points have been cited:

- Observation is carefully planned, systematic, and perceptive. Observers know what they are looking for and what is irrelevant in a situation. They are not distracted by the dramatic or spectacular.
• Observers are aware of the wholeness of what is observed. Although they are alert to significant details, they know that the whole is often greater than the sum of its parts.

• Observers are objective. They recognize their likely biases, and they strive to eliminate their influence upon what they see and report.

Attention was made in choosing such a tool for collecting data in the present study. The main advantage of observation according to Kothri (1990), for instance, is that subjective bias is eliminated, if observation is done accurately. Secondly, the information obtained under this method relates to what is currently happening; it is not complicated by either the past behaviour or future intentions or attitudes. Thirdly, this method is independent of respondents' willingness to respond and as such is relatively less demanding of active cooperation on the part of respondents as happens to be the case in the interview or the questionnaire method.

In accordance with the above arguments and justifications for using 'observation' as a method of data collection, a schedule of observation was developed with special attention paid to ensure consistency in such a tool. The schedule was designed to account for all necessary and very much related aspects to the present study (See appendix 24). Several reading classes were attended by the researcher as a non-participant observer in both group A (first year class) and group B (second year class) with their full numbers as usual as any other natural classes. The recordings of the observations of those classes were done immediately after attending each observed class to insure the inclusion of every detail a fresh. Later on, and as part of data analysis these observation schedule forms were analyzed and discussed.

3.4.5. Questionnaire:

Experts in both educational and behaviuorial research like Best & Kahn (1993), Koul (1997), Gupta (2001), Kothri (1990) among others, acknowledged the importance, practice, and popularity of questionnaires as tools in collecting data in social and educational studies. In their
A discussion of questionnaire as a tool in educational research. Best and Kahn (1993:229-230) recognize "the general category of inquiry forms includes data-gathering instructions through which respondents answer questions or respond to statements in writing". They further make a distinction between two tools that look alike with some slight differences. A questionnaire, according to them, is used when factual information is desired. When opinions rather than facts are preferred, an opinionnaire or attitude scale is utilized.

However, when information cannot be obtained through other tools like school reports or census then questionnaire is to be adopted. When designing a questionnaire, much care should be given to its significance which should be carefully and clearly stated. In defining questionnaire, Kothri (1990:124) asserts that a "questionnaire consists of a number of questions printed or typed in a definite order on a form or a set of forms. The questionnaire is mailed to respondents who are expected to read and understand the questions and write down reply in the space meant for the purpose in questionnaire itself. The respondents have to answer the questionnaires on their own." Additionally, Koul (1996:146) in his definition of questionnaire considers the matters such a device can be utilized for. It can be of the objective of gathering information on either of psychological, social, or educational topic(s) that can be sent or given to individuals or group of individuals, as in the case of this study, with the aim of attaining data related to some problems under research.

Nevertheless, questionnaire is a popular tool of gathering all kinds of data in research. The questionnaires can be classified into two types in terms of the nature of the questions constructed in a questionnaire. Questions that require only a short or check responses are called 'closed questionnaire', whereas those needed to be answered in 'open' category of response are known as 'open questionnaires'.

There are two types of questionnaires were designed to elicit information for this study. One was designed for the students (SQ) and the other one was designed for the teachers (TQ).
3.5. Procedure of Collecting the Data:

In the present study, the data collection was done in a period of well over a month in the Department of English, College of Education, at Ataq, Shabwa; Aden University. Three cloze tests, three comprehension tests, one composition test for each group were administered, along with two questionnaires that were filled by both learners and faculty. Finally, some teaching classes of reading were attended by the researcher, as mentioned earlier in this chapter. However, each test took 60 minutes to be done, and these tests were carried out in a frequency of every other day and sometimes a gap of three to four days was left due to the full class schedule of the two levels in the department and the two-day weekend.

Firstly, three comprehension tests were given to the subjects of both groups. Clearly stated instructions were the first thing to begin in each session of these tests so that students can begin confident and natural and their responses well done.

Secondly, three cloze tests were administered in the same manner, as far as frequency is concerned. Those three cloze tests were designed and developed from the same texts tested for the three comprehension tests of each group. Regarding time, cloze tests sessions were approximately done over an 80 minutes period for each.

Thirdly, composition tests were administered for the two groups' subjects. In each test the students were asked to write a free composition on an assigned topic, as mentioned previously in (3.5.3). Each session of these composition tests was done in an 80 minutes period.

Fourthly, on completing these tests, a questionnaire was distributed among the whole students of the department (TQ), who were present at the time of carrying out such a tool. About 120 copies of the students' questionnaire form were circulated by hand to the students, and instructions were clearly explained in the classes. Unfortunately, only 103 copies had been received back from the students, 15 of these forms did not satisfy the complete required information, i.e., some of these forms were either returned completely without response or partially answered.
However, neither of which was acceptable, so these 15 forms were totally exempted from the analysis.

Furthermore, another questionnaire is the (TQ), which was designed and distributed among the teachers of the department to elicit information about their attitude towards teaching English and teaching reading in particular, teachers, qualifications teaching experience, time allocated for reading, their evaluation for their students performance, availability of teaching/learning materials, intelligibility of the library and some other related matters. Six of the teachers, who are the whole number at the time of the data collection, took part in the questionnaire. Six forms were circulated among the teachers and all of them returned back on time.

Finally, observations of several classes of teaching reading were conducted by the researcher. Those classes were attended to get more information about some issues related to teaching/learning process of reading that could not fully gained through tests and questionnaire. The researcher attended as a non-participant observer and observed the actual situations of teaching and learning reading in classroom. Needless to say, after each session of observation an observation schedule form (see appendix 24) is filed by the researcher to ensure that all information gained are recorded afresh.

Having discussed the methodology of this research, let us have a detailed discussion on the data analysis of this study.