METHODOLOGICAL FRAMEWORK
Teaching and testing are not separate entities but go hand-in-hand, for these are complementary with each defining the shape of the other. Teaching has always been a process of helping others to discover new ideas and new ways of organizing that which they learned. Whether this process took place through systemic teaching and testing, or whether it was through a discovery approach, testing was, and remains, an integral part of teaching. Just as there is no best teaching method, there is no best method of testing.

One of the earliest researchers on language testing, Davies (1968:5) says that a “good test is an obedient servant since it follows and apes the teaching.” We find ourselves in complete agreement because the relationship between teaching and testing is surely that of partnership in the context of complementarily. It is true that there may be occasions when teaching is good and appropriate and testing is not. In such a situation, we are sure to be confronted with a harmful backwash. This effect seems to be a situation of the kind Davies had in mind when he saw testing performing the role of servant to teaching. This would be oversimplification of the
problem because there may be occasions when teaching is poor or inappropriate but testing is able to exert a beneficial influence. We cannot expect testing only to follow teaching. What we should demand of it, however, is that it should be supportive of good teaching and, where necessary, exerts a corrective influence on bad teaching. If testing always had a beneficial backwash on teaching, it would have a much better reputation as a tool for teachers to discover what they have been able to teach and to what extent that has worked with the students.

4.2 The other reason for mistrusting tests is that they very often fail to measure accurately whatever it is that they are intended to measure. Teachers know this. The true abilities of students are not always reflected in the test scores that they obtain. To a certain extent this is inevitable. Language abilities are not easy to measure and we cannot expect a level of accuracy comparable to those of measurements in the physical sciences. However, it would not be out of place to expect greater accuracy than is frequently achieved.

4.3 Why makes test inaccurate is the fact that there are two main sources of inaccuracy in it quite often. One concerns test content and techniques. For example if the problem is one of how well someone can
write, then a multiple choice test would in no way help us get a really accurate measure of his/her ability. A lot of research effort has gone into this already and yet one cannot say with cent per cent confidence that it has been successful. There are still numerous loopholes that need to be plugged.

The second source of inaccuracy is lack of reliability. A test is reliable if and only if it measures consistently what it is intended to measure. A reliable test can used with confidence that the testees will get more or less the same score, irrespective of the fact that they happen to take it on one particular day or on the next. An unreliable test is thus marked for the different scores that it is likely to give even if it were to be administered a second time on the same day.

4.4 Unreliability has two origins: features of the test itself and the way it is scored. Lack of reliability of a test that poses problems arising out of the features throws up results wherein a testee shows the tendency to perform significantly differently on different occasions on which he/she takes the best. Performances might differ greatly across different occasions. Consequently, even if the scoring of a testee’s performance on the test is perfectly accurate (that is, the scorers do not make any mistakes), the testee
will nevertheless obtain a markedly different score, depending on when he/she actually sat the test despite the fact that there has been no change in the ability which the test is meant to measure. The examples of all possible features of a test which might make it unreliable are: unclear instructions, ambiguous questions, items that result in guessing on the part of the testees. While it is not possible entirely to eliminate such differences in behaviour from one test administration to another, there exist principles of test construction which can help the teacher to minimize these to whatever extent possible.

The way a test is scored is also important. Equivalent test performances are accorded significantly different scores. For example, the same composition may be given very different scores by different markers or even by the same marker on different occasions. This happens when there is no scoring key on which the markers are in full agreement. Or, even when there is no bench marking as in competitive examinations, wherein the markers are given a number of scripts to mark, especially on an essay test, and the marked scripts are then evaluated by a moderator to ensure unfamiliarity across markers. The moderator then issues instructions to individual markers with regard to their evaluation.
4.5 We shall focus our attention now on test specifications. These are an official statement of what the test seeks and how it tests it. It is a description of the test, written beforehand, which includes such information as what the purpose of the test is, what will be covered by the test, how it will be tested etc. Test specifications are written before the test, and then the test items are drafted on the basis of these. Once the drafting of the test completed, these specifications are used to ensure whether the test items match so draft match with them, preferably by someone not directly connected with the writing of the test. This is normally done for large standardized tests, but it is also useful practice for individual teachers making tests for their own students. There is, however, the problem of confidentiality in such a situation. Test specifications can help the teacher clarify whether the test covers what it is intended to cover in the way it is supposed to cover it. We could not have contacted each and every student and ask for his/her opinions individually due to constraints of time, and the availability of students. Television, internet and newspapers were the media which would have been of great help to us in conducting this study but the idea of using them had to be abandoned because these are very costly and require a long period of time. So we decided to adopt a random survey method.
4.6 Since the survey reported here is meant to take into account the standards of English language among the students enrolled in our educational institutions in Gujarat, we felt the need to adopt suitable measures to conduct the survey. The researcher had referred to the DAI (Dissertation Abstracts International) to ensure that the topic that had chosen was not a repetition of any other research work. Raviya (2002), like others before him, offers us insight into the traditional testing practices in vogue. Therefore, we decided to avoid duplication. We took these studies as the point of reference from which to proceed. Accordingly, two devices, i.e. (1) a test package of the kind presented in Cyril Weir (1985), and (2) an opinionnaire validated by J Charles Alderson for students were used with certain adaptations.

4.6.1 The test package contained three tests - listening test, guided note taking test and reading comprehension test. The time allotted to each test was 25 minutes, 50 minutes and 75 minutes respectively. In all the three tests the texts were in written form. The test package aimed at testing students’ language skills i.e. listening, speaking, reading, writing, and study skills i.e. reference skills, note taking, note making, summarizing, time management etc.
4.6.1.1 The first test in the test package, which was appended as Annexure 3, was designed to test respondents' ability to understand spoken English. They had one task to do in approximately 25 minutes. They heard a short tape recording once only. The passage was recorded in the voice of the supervisor of this research work in Gujarat recording studio at Anand. After the listening they were given a cloze test which contained 20 items.

4.6.1.2 The second test item in the test package is meant to test students' ability to understand spoken English and note taking skill. They had to make notes and use them to answer a number of questions. They had two tasks to do in approximately 50 minutes. They heard a tape recording of a short lecture only once. This was done so as to create a lecture situation. No lecture is normally repeated because it is not a drill. A written outline of the main points of the lecture was printed in the answer booklet to help them to follow what the speaker was saying. This lecture outline consisted of important statements from the passage, each followed by questions. While listening to the lecture they had to make notes so that they would have time to go through these notes after the lecture and use these to write answers. They had to write a summary of parts of the lectures, using the lecture outline and their notes and answers.
4.6.1.3 The third test item in the test package is test of students’ ability to read in English and to write in English about what they have read. They had two tasks to do in 75 minutes. To test their reading comprehension, note making and organizational skills, a passage of approximately 650 words was given them to read. They had to write short answers to 15 questions after reading the passage. There are 10 ‘wh’ type questions, 4 ‘explain’ type questions and 1 is ‘summary’ type question. To help them to do this, they were required to take brief notes while reading the passage so that they could use these to make a summary for themselves.

4.7 The second tool was an opinionnaire. An opinionnaire was prepared to know students’ opinions about their course in English and thus to design a remedial programme for them. 210 students of third year B.Sc. from seven science colleges of Gujarat were selected, but not in terms of intelligence because the opinions and observations of the average and poor students are also equally valuable.

4.8 In terms of its purpose, a test generally falls into one of five broad categories: placement, progress, achievement, proficiency and diagnostic. It is important before starting to write a test to know which of these broad purposes the test has. Our purpose is to test achievement of
students in English in the sciences at the undergraduate level in Gujarat. Achievement tests are directly related to language courses, their purpose being to establish how successful individual students, groups of students, or the courses themselves have been in achieving objectives. There are two kinds of achievement tests: final achievement tests and progress achievement tests. Final achievement tests are those administered at the end of a course of study. They may be written and administered by ministries of education, official examining boards, or by members of teaching institutions. Clearly the content of these tests must be related to the courses with which they are concerned, but the nature of this relationship is a matter of disagreement among language testers. Progress achievement tests, as their name suggests, are intended to measure the progress that students are making. Since 'progress' is towards the achievement of course objectives, these tests too should relate to objectives.

4.9 Achievement tests may be used in four ways to make decisions about individuals: as a survey of attainment in a content area, as a diagnostic instrument to identify individual strengths and weakness in a content area, as an indicator of readiness to determine if an individual has attained prerequisites to continue study in a given content area, and as a performance test to estimate the degree of learning of a body content.
The students fall into the age group of 20 to 21 years. It has long been felt that there is a strong need to bridge the gap that exists between the communicative ability in English of the students in the sciences with reference to the expected level of communication in their fields. There are college teachers of English today who admit (though not in public) that they have stopped bothering about their students. The common experience of teaching English in technical colleges is that 80% students are not motivated enough to study and education out the collegiate level just another pastime. The English classroom at the undergraduate level, therefore, becomes a monologue, with the teacher as the performer and the students as passive observers. It has been observed that the rest, i.e. the remaining 20% of the students is highly motivated, bright and articulate. They complain of lack of challenge in the classes. The linguistic base with which most science students enter college is nothing more than a mere fragmentary knowledge of the syntax of English language, which is totally inadequate in meeting their professional demand. On considering their sociolinguistic profile we see that they come from rural areas, start learning English language learning from fifth standard and this they do in environment wherein learning is completely missing. Even with a background of this kind, they find themselves thrown into a situation where they learn all the subjects in English and are expected to understand these
and express themselves in reasonably correct English. Of course, what is ‘reasonably correct English’ for a teacher may not be so far another, which is what makes this phrase a misnomer.

4.11 The English courses they go through at this level hardly teach them to think, understand and express independently even though they clear those courses satisfactorily. The problem becomes more serious when they face the language of the teachers of other subjects. Somehow they manage to transfer the information but the language they use is not up to the mark. The result is that if on the one hand, they are given exposure to English as a medium of instruction by a group of teachers whose English language proficiency may not be of the level expected, there is on the other hand, a situation wherein students, despite their poor competence, manage to get good marks in the technical subjects. They eventually lose interest and do not feel the need to do anything to improve their level of competence in English. This makes it a great challenge for English teachers to motivate the learners to develop an interest and ability of the kind that would enable them to handle adequately the skills of English language.

4.12 Then there is the problem of relating the topic to the learner’s needs and interests. An average college student learning English may have
little ability, but he/she has a more mature mind and, thus, the facility of a fair amount of sophistication in the use of language. Many of the classroom techniques designed to teach language skills at the college level, such as oral drilling have been tried out often but with rather poor results. For a mature learner, language learning should be more challenging than just carrying out repetitive tasks. The breaking down of the skills into listening, speaking, reading and writing is only a broad and rough kind of classification. There may be two main parameters in the design of courses in the light of English in India being treated as a library language: the register to be used and the skills to be concentrated upon. The student does not need to command any register except the one he/she is concerned with and, therefore, he/she need only be exposed to the vocabulary and structures of the register. This principle is applicable in designing courses in English for the students of Agriculture, Science and Technology etc. While a multi-skill approach may be preferable at the elementary level, it may be desirable later to concentrate on a particular set of skill that the learner needs.

4.13 In the institute where this researcher teaches English, most students come with rural background, with the vernacular (Gujarati) as the medium of instruction up to the +2 stage, i.e. the Higher Secondary level.
The students in general are not interested in the courses in English and complain that they have a heavy load of technical subjects, which are more important for them in shaping their future career. The very idea of a general English course of the kind they have already studied for several years at school, without any apparent gain, is enough to make them indifferent towards the courses.

4.14 It is necessary to find out whether or not students seeking admission to formal education programmes possess adequate language and study skills to be able to follow the course offered as a part of the programme and to suggest remedial measures to correct inadequacies, if any.

4.15 The objectives of this research are: To measure the level of achievement reached by a testee as a learning outcome. For this, it is essential:

(a) to find out the extent of mastery of study skills attained;
(b) to place the level reached in terms of the mastery of reference skills;
(c) to find out the extent of possible level of competence reached;
(d) to find out whether or not the testee is good with regard to time management, and if 'yes', to what extent.
4.16 We have made an attempt to test the following hypotheses:

The candidates are not, at least in majority of cases, able to

(a) possess adequate written English for formal academic writing tasks;
(b) understand spoken English for listening lectures and discussion;
(c) understand written English for reading text books and other 
   sources of information both intensively and extensively;
(d) have adequate skill of oral expression for taking part in academic 
   discussions and presenting papers.

4.17 We hope that to be able to

(a) help design course and a testing package in English appropriate for 
   the students in the sciences and
(b) also help make the teaching and learning of English for educational 
   purposes more meaningful on the basis of this research effort.