Exploring the Relationship between Classroom Language Assessment Literacy and Practices Using

A Short-term Teacher Education Course

A dissertation submitted to the University of Hyderabad for the partial fulfilment of the requirements for the degree of

Doctor of Philosophy

in

English

by

Santosh Kumar Mahapatra

Supervisor:

Prof. Pingali Sailaja
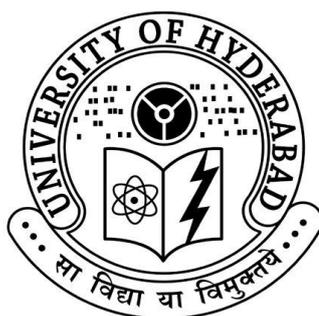
Department of English

School of Humanities

University of Hyderabad

Hyderabad, India- 500046

March 2015

Department of English

School of Humanities

University of Hyderabad

This is to certify that I, Santosh Kumar Mahapatra, have carried out the research embodied in the present dissertation titled *Exploring the Relationship between Classroom Language Assessment Literacy and Practices Using a Short-term Teacher Education Course*, for the full period prescribed under the Ph.D. ordinances of the University.

I declare that to the best of my knowledge no part of this dissertation was earlier submitted for the award of research degree to any university.
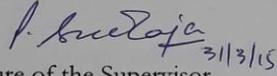
Signature of the Supervisor

Professor Pingali Sailaja
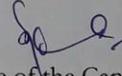
Signature of the Candidate

Name: Santosh Kumar Mahapatra

Registration No: 11HEPH03

Head

Department of English

Dean

School of Humanities

ACKNOWLEDGEMENTS

ABSTRACT

A mixed methods study, this doctoral research explores the relationship between the classroom language assessment literacy (CLAL) and assessment practices of English teachers. It is based on the premise that if English teachers are properly oriented in assessment of language ability in classroom, the quality of classroom assessments will improve and in turn, teachers will be in a position to carry out "assessment for learning" effectively. The study is designed considering the assessment policies and practices in India as a whole. The intervention takes place in three schools— a CBSE, a State-Board and an ICSE school— and six secondary school English teachers— two each from the above-mentioned schools— in Andhra Pradesh constituted the main participants. This research is significant in the sense that though the new assessment policy in schools demands that teachers carry out classroom assessments, and attaches a lot of importance to those, very little has been done to equip teachers with adequate skills and knowledge to enable them to meet the challenges of classroom assessment. This study tries to draw attention towards this problem and suggests a possible solution.

For its theoretical framework, the study draws on classroom assessment, assessment literacy, second language teacher education, teacher development, teacher beliefs and teacher knowledge. The framework is built after studying and tracing the existing gaps in the research literature in areas relevant to the study.

The study is conducted in three major phases: pre-intervention, intervention and post-intervention. In the first phase, the researcher collects information about the CLAL of secondary school English teachers practicing in CBSE, state-board and ICSE schools through a state-wise survey. Then two teachers each from a CBSE, a state-board and an ICSE school are selected and information about their CLAL, beliefs about assessment and assessment practices is obtained. A teacher development (TD) programme based on the above-mentioned information is designed with a view to developing teachers' CLAL. In the second phase, the programme is imparted separately to the three pairs of teachers for 18 hours. Records of teachers' experience and the researcher's observations are maintained during the intervention. In the last phase, once again, information is collected about the CLAL and assessment practices of the teachers in a similar manner as done in the first stage. Data are analysed both qualitatively and quantitatively, triangulated and

meta-inferences are made after a few cross-case analyses. Finally, an effort is made to generalize and discuss the findings, state the implications of the study for policy makers, teacher education, curriculum designers, experts in language assessment and teacher educators, point out the limitations of the project and suggest directions for further research in the area of the study.

The findings of the study suggest that there is a relationship between CLAL and assessment practices of teachers. But a lot of factors like the assessment policy at curricular level, institutional policy about assessment, motivation of teachers to utilize assessment for pedagogic purposes and their English language ability are found to have impact on teachers' assessment practices. It is also found that a need-based TD programme can, to a great extent, generate the desired impact on teachers' assessment practices.

TABLE OF CONTENTS

CHAPTER 5 IMPACT OF THE THER DEVELOPMENT PROGRAMME ON TEACHERS' CLASSROOM LANGUAGE ASSESSMENT LITERACY

CHAPTER 6 EXAMINATION OF THE RELATIONSHIP BETWEEN TEACHERS' CLASSROOM LANGUAGE ASSESSMENT LITERACY AND ABILITY TO CARRY OUT CLASSROOM ASSESSMENTS

CHAPTER 7 DISCUSSION AND CONCLUSION

LIST OF TABLES

LIST OF FIGURES

DEFINITION OF TERMS

There are a few terms, which are used in the dissertation to convey certain meanings. They are defined below.

Assessment is used as an umbrella term and includes testing and evaluation of formal and informal nature under its purview.

Assessment literacy refers to the knowledge about assessment, the ability to carry out assessments and appropriate beliefs about assessment.

Classroom assessment is assessment that is carried out in the classroom by the teacher for assessing the progress of their learners, and is also called formative or teacher-based or internal assessment.

Classroom language assessment literacy is the assessment literacy with regard to classroom assessment of language ability.

Experts in language assessment are teachers at university level with research and teaching experience in relation to language assessment.

Language assessment literacy is assessment literacy with regard to assessment of language ability.

Principles of assessment comprise principles of Validity, Reliability, Authenticity, Practicality and Washback as applied to language assessment.

Teacher assessment ability refers to the ability of the teacher to carry out assessments.

Teachers' assessment practices include the practices of planning, designing, conducting and evaluating assessments, offering feedback to students and analysing and using results of assessment properly.

CHAPTER 1

INTRODUCTION

1.1 Aim of the Study

The study proposed to explore the relationship between classroom language assessment literacy (CLAL) and classroom assessment (CA) practices of secondary school English teachers. It was based on the premise that assessment is an integral part of teaching, and it is important that teachers are convinced about this. But in a country like India, where teachers do not have sufficient knowledge about assessment, and very little effort has been made to educate teachers in this direction, there is an urgent need to develop and conduct small scale need-based in-service teacher education programmes for them (National Council for Teacher Education, 2010, p. 39). The study was built on the belief that test-centred nature of teaching can be changed into assessment-propelled learning if teachers are trained in language assessment.

1.2 Background to the Study

In India, the learning of English at school level has not been an entirely burden-free experience. The emphasis on learner-autonomy and learner-friendliness in policy documents like *National Curriculum Framework* (NCF) *2005*, *Position Paper- National Focus Group on Teaching of English* (National Council of Educational Research and Training, 2006), *National Curriculum Framework for Teacher Education* (NCTE, 2010), etc. has not been transferred to classrooms (NCTE, 2010, p. 4). Moreover, practices in ELT across the country are still heavily dependent on textbooks and driven by examinations. It will take some time before awareness about language proficiency spreads across most language classrooms and societies in the country. Though there have been changes in the attitudes towards teaching and learning of English, these are too slow to keep pace with the growing demands for English in the country.

The quality of ELE (English Language Education) in India is marred by the use of traditional methodology by the teacher. Due to the "problems of systemic feasibility and preparedness" (NCERT, 2006, p. 1), it has been difficult to get well-trained teachers to teach English in schools. The lack of training gets reflected in their teaching. Inadequate

training could be a strong reason why effective innovations and research in the field have not been implemented in classrooms. For example, not many teachers are aware of how to make use of the existing language awareness of bilingual/multilingual learners for improving their proficiency in English. However, there are schools in which English teachers are properly trained and linguistically competent. These teachers are better-equipped to enable their students to become effective users of the language. However, the percentage of students taught by such teachers is very small when compared to the huge percentage of students, who have very little access to quality teaching.

The lack of proper training for teachers working at secondary level has been a concern (NCTE, 2010). It has left teachers only with their naturally developed notions about teaching and learning. In most cases, such notions have their roots in the classroom teaching of their own teachers. Then, there are issues like adherence to lecturing and teacher-centred ways; overemphasis on mastering textbook content and scoring good marks in examinations; too much dependence on textbooks for materials; an uninformed sense of language proficiency and assessment, etc., which are often associated with practising teachers. These claims are mostly impressionistic, but very few can deny that there is some amount of truth in these claims. Here, the concern is that the aforesaid gaps in teacher knowledge can handicap a teacher, to a great extent, especially, while noticing and addressing students' language related problems and paying attention to individual language needs.

1.2.1 Assessment of English in Schools

The unwarranted encouragement given to high scorers mostly by parents and sometimes by teachers has not helped the field of language assessment. As pointed out in the *Position Paper- National Focus Group on Teaching of English* (NCERT, 2006), examination scores do not always reflect the student's original ability to use language in communicative situations. But teachers cannot be blamed for failing to design and use sound classroom assessments and gather information about students' progress and problem areas because very little professional support is provided to teachers in this direction. As suggested by the *Position Paper, National Focus Group on Examination Reforms* (NCERT, 2006), entrusting teachers with all the responsibilities related to classroom assessment, thus, is unfair.

Assessment has been a necessary and yet, highly examination-dominated practice in India. It has often led to high level of anxiety, rote memorization, high rates of failure,

unhealthy competition, suicides, etc. on the part of students as reported by the NCF (NCERT, 2005), *Position Paper- National Focus Group on Curriculum, Syllabus and Textbooks* (NCERT, 2006) and NCFTE (NCTE, 2010). What is practiced in schools in the name of assessment can be easily "associated with examination, stress and anxiety" (NCERT, 2005, p. 71); and such practices "are highly inadequate and do not provide a complete picture of an individual's abilities or progress towards fulfilling the aims of education" (NCERT, 2005, p. 72).

English is perhaps the most feared subject among secondary school students, especially for those who are in regional-medium schools. In these schools, the rate of failure in English in board examinations is more than in any other subject. Insufficient exposure to the language; improper teaching strategies; lack of variety in teaching materials; problems with the construction of the examination question paper, etc, could be some of the reasons for students' dismal performance. The examination results often show what students cannot do with the language. Board examinations promote an unhealthy race for scoring marks for some, and for others, a matter of passing and failing. The quality of examination question papers, their effectiveness in measuring student performance, the accountability of test designers, the expertise involved in the process, the gaps between what is taught and tested and what is aimed at being tested and actually tested, etc. are some of the issues that need more attention at policy level.

1.2.2 Blame on Teachers

For some reason, teachers get a healthy share of the blame for the negative effects of assessment. In the NCF 2005, it stated that "[o]ften children's learning is restricted as teachers do not accept their answers if they are different from what are presented in the guidebooks" (p. 74). It is ironical because most of the times, the teacher does not have control over the content of the test and what should be marked as the/a correct answer. A teacher has to work under a fixed framework, about which, he/she does not have much say. Policy makers seem to ignore this factor. It becomes evident when one looks at NCF 2005, in which, it is clearly stated that teachers should spend more time on designing question papers, giving open-book tests and providing constructive feedback to students (pp. 71-76). Even in the *Position Paper- National Focus Group on Curriculum, Syllabus and Textbooks* (NCERT, 2006), it is mentioned that teachers adhere to repetition, drilling and other rote-memorization techniques to make students score well in the examinations. But the lack of pre-service training in teaching methodology is hardly

discussed in policy documents. In an effort to counterbalance this blame on teachers to cover up the loopholes in the system, a brief discussion on the need for professional development of teachers is presented below.

1.2.3 Examination Reforms and CCE

On the basis of recommendations made by several committees and commissions, especially NCF (2005), the government decided to change examination practices. What was introduced as a platform for reform was Continuous Comprehensive Evaluation (CCE). The aim was to lessen the burden on children, make the assessment process inclusive and systematic, provide more flexibility to the teacher for teaching, and diagnose students' problems with learning. According to the recent changes in assessment policy, weighting is given to the internal assessments carried out by school teachers for their own students. This decision seems to be driven by the belief that teaching can be improved by changing the examination pattern. However, a similar approach has already been proved ineffective by Cheng (2005) who, in a large-scale study conducted in Hong Kong, found that "changing the examination does not change the degree of emphasis on examination nor does it necessarily change teachers' methods of teaching in any fundamental way, if teacher education and professional development are not involved" (p. 251). Although India and Hong Kong are socio-culturally different from each other and have dissimilar educational systems, Cheng's findings certainly cannot be ignored. The findings also imply that there could be a possible relationship between teachers' professional development or training in assessment and their belief about and attitude towards assessment and ability to assess.

There remains much to be resolved before the aims of CCE are achieved within the realities of English classrooms in schools. An English teacher should know how to use "a variety of assessment tools and techniques" (NCERT, 2006, p. 115), analyse and interpret student performance, follow five basic principles of language assessment while designing assessment tasks, keep track of students' language learning process and development of proficiency, assess what learners know, etc. to be able to do justice to CCE. Very few will disagree that the list is a long one, and it puts unrealistic demands on the English teacher precisely because the basic requirements for an English teacher - like proficiency in the language, training in language teaching methodology and materials design, etc. - still remain unfulfilled in most cases. This change in assessment policy can be interpreted as "a political response to people's aspirations rather than an academic or

feasibility issue" (NCERT, 2006, p. v). It may not be a reform that will result in better teaching and learning. The NCF (2005) could foresee this problem and warned against any change in policy before getting teachers trained and ready for implementing the changes.

1.2.4 Assessment Training in English Teacher Preparation/Development Programmes

All the stakeholders in a formal teaching-learning scenario, especially curriculum designers, policy makers, test designers and above all, teachers need to have some amount of assessment literacy. Assessment literacy has been defined as an awareness of principles for designing thorough assessments (Stiggins, 2002; Popham, 2004). But, the amount of assessment literacy required for a curriculum designer or a textbook writer or a test designer is expected to be more than that of a teacher and varies from one educational context to another. It may be necessary for teachers to have the skills to apply the principles of assessment, analyse and report student performance and suggest necessary changes in approaches to teaching, instructional materials and assessment policy. Though there is no universally-accepted definition of assessment literacy, it generally refers to "skills", "knowledge" and "principles" required to design, conduct and report assessments (Davies, 2008). Going by this definition, it becomes obligatory for teachers to be assessment literate so that they can assess their students efficiently. Even when they do not have the responsibility of assessing their students, their knowledge about assessment can help them to develop a healthy attitude towards teaching. But most teacher education systems in the world have not shown much interest in developing assessment literacy of English teachers. This has been claimed by many researchers including Schafer (1993), Popham (2006), Stiggins (2007), Rogier (2009) and Coombe, Troudi and Al-Hamly (2012). And the situation is not very different in India.

Except for B. Ed. programmes offered at English and Foreign Languages University, and H. M. Patel Institute of English Training and Research, there is no exclusive formal pre-service teacher preparation programme of good quality for English teachers in the country. Moreover, language assessment does not get satisfactory attention even in these programmes. In other B. Ed. programmes across the country, a trainee (from any academic background) can opt for English Teaching Methodology along with methodology for teaching a content subject and can be officially eligible to teach English. Apart from the ones offered by famous institutions like Regional Institute of Education,

5

Central Institute of Education (Delhi University), etc., very few B. Ed. programmes across the country offer active training in English Language Teaching Methodology. Even fewer programmes offer help to English teachers with regard to developing theoretically and practically sound assessments. So a large percentage of English teachers starts and continues teaching in schools without having minimum knowledge about language assessment. They get some in-service training provided by DIETs, SCERTs, ELTIs, some universities, etc. These programmes are 'sporadic' in nature and have "limited relevance to needs" (Padwad, 2011, p. 11) of teachers. The number of programmes exclusively devoted towards developing teachers' ability to carry out classroom assessments effectively is far fewer than what is actually required.

1.3 Significance of the Present Study

Teacher development in language assessment is central to the success of teacher-based assessment followed in CBSE and schools run by many State Boards across the country. Since most teacher preparation programmes in India do not offer adequate training to prospective English teachers in language assessment, it is necessary to provide in-service training to teachers. This study makes an attempt to show that proper training in language assessment leads to enhanced level of classroom language assessment literacy (CLAL) and an enhanced CLAL level can ensure better assessment practices. The findings of the study are very relevant to the present assessment scenario and may be expected to invite more research attention to the area under scrutiny. The study is expected to provide some impetus to the efforts towards educating English teachers in important aspects of language assessment.

1.4 Hypotheses

Against the background discussed in the previous sections of this chapter, the current study was based on the following hypotheses, which were specific to the context of Andhra Pradesh, where the study was carried out:

- There is a relationship between teachers' level of assessment literacy and language assessment practices.

- A need-based TD programme in CA of language ability will lead to change in their assessment practices.

1.5 Research Questions

The study tried to address the following research questions:

- What is the average CLAL level of secondary school English teachers in the state?

- How is teachers' CLAL related to their assessment practices?

- What impact does a short TD programme in CA of language ability have on the teachers' ability to design classroom assessment tasks and assessment criteria, provide feedback, and on the teachers' beliefs about assessment?

- How do teachers respond to the programme?

1.6 Organization of the Dissertation

The dissertation has a total of seven chapters. This division was done keeping in mind some general academic norms followed for organizing doctoral dissertations, a logic of systematic presentation and the readability factor.

The first chapter describes the aim, background and significance of the study along with the hypotheses on which the study was based and the research questions it addressed. The next two chapters present a review of relevant research literature related to the topic. The methodological design used for the study is discussed in the fourth chapter. The next two chapters contain presentation and analysis of the data collected for the study. The concluding chapter of the dissertation focuses on findings and discussion of findings, implications, suggestions for future researchers and limitations of the study.

1.7 Conclusion

This chapter has described the aim of the study and the background to the study, as also the problem under scrutiny, the hypotheses and research questions for the study, and organization of the dissertation. The next two chapters deal with a review of research relevant to the current study.

CHAPTER 2

CLASSROOM ASSESSMENT AND TEACHER ASSESSMENT LITERACY

2.1 Introduction

This chapter is divided into two main sections to give adequate importance and space to the areas under review. The first of the two sections focuses on Classroom Assessment (CA) and includes discussions on salient features of CA and Formative Assessment (FA). The next one is related to Teacher Assessment Literacy. Various definitions of Assessment Literacy and their relevance to English Language Education (ELE) are discussed and an eclectic framework for Assessment Literacy is developed. This framework takes into consideration teacher-based assessment of learners' English language ability in the classroom and the practice of the same at the secondary school level in India. However, this framework is not rigid and thus, can be relevant to other second language situations of a similar kind.

2.2 What is Classroom Assessment?

CA is a learner- and learning-friendly approach to assessment. The definitions of CA have not changed much over the years. According to Cross and Angelo (1988), it is a means for teachers to collect information about "the level and quality of student learning" (p. 2) and involves merging of "assessment techniques" with "teaching tips" (p. 5). The features of CA included in this definition are echoed later by Popham (1995) who defines it as a formal practice by the teacher to know about their students' strengths, weaknesses and progress, give them grades and find out the effectiveness of the course. Cizek (1997) and Airasian (1997) also share similar views.

In a huge study comprising 500 CAs collected across 10 years, Black and William (1998) find that CA is a student-centred practice which provides descriptive rather than evaluative feedback on teaching and learning. The student-participation factor is further confirmed by Brookhart's (2001) study. The study indicates that students can take responsibility for their learning and do self-assessment at regular intervals if the teacher uses CA properly. The Assessment Reform Group (2002) of Britain adds an extra

dimension to the existing definitions of CA by asserting that apart from being highly motivating, it promotes learner-involvement, provides constructive feedback on learning and helps in curricular goal-setting among other things.

An obvious characteristic, the concept of "local assessment" is explicitly associated with CA in the definition provided by Leung (2005). According to him, it is a kind of "non-standardized local assessment carried out by teachers in the classroom" (p. 871). In this case, the word "local" indicates that every classroom is a unique one and thus, the teacher who handles the classroom knows which kinds of assessments serve the purpose of ensuring and enhancing learning in their classroom context.

From the above definitions, it is obvious that CA can be defined in many ways and has several dimensions. The dichotomy between formative and summative, formal and informal or standardized and non-standardized continues to be part of the debates related to classroom-based assessments. A multimodal definition can accommodate a lot of these aspects: "CA is the planned collection of information about the outcomes of teaching on student learning." (Shermis & Di Vesta, 2011, p. 2). This definition has many layers to it. However, we do not need an over-inclusive definition. A definition that fits the requirements of a particular educational context or school or class or even group of students can be useful. The implication is that what works for promoting learning in a particular context can go into the making of the CA in that context.

2.3 Characteristics of Classroom Assessment

The main characteristics of CA can be found in its definitions mentioned in the previous section. But a focused discussion can throw further light on its characteristics in some detail. As pointed out by Angelo and Cross (1993), CA is "learner-centred, teacher-directed, mutually beneficial, formative, context-specific, on-going, and firmly rooted in good practice" (p. 4). Authenticity can also be added to this list. Some of these important characteristics are discussed briefly in the following sections.

2.3.1 Learner-centredness

Most researchers working on CA agree that learners are central to this form of assessment (Rodriguez, 2004; Stiggins, 2004; Frey & Schmitt, 2007; Obeg, 2009; Stoynoff, 2012, etc.). This implies that the focus is on maximizing learning through learner-friendly and yet engaging and productive assessment practices. Every individual

student gets attention from the teacher. The teacher utilizes a variety of assessment methods and aims to elicit the best performance from each student. In the process, students learn to become self-dependent and be responsible for their own learning (Earl, 2012).

2.3.2 Teacher-directedness

Though learners hold the centre-stage in CAs, it is teachers who lead the process (Stiggins & Conklin, 1992). Several decisions are taken by the teacher even though students are involved in the process. From choosing assessment methods to grading and giving feedback, the teacher has many things to take care of. However, this entire process does not essentially hinder learner autonomy because learners' interests top the list of priorities in CAs. The fact is that on most occasions, the teacher knows his/her students better and is aware of their learning preferences and the developmental process. And this puts the teacher in a better position than others when it comes to designing suitable assessments for his/her own students. Especially in countries like India, teacher-developed assessments can be of great help considering the fact that classrooms differ from each other in terms of socio-cultural practices, mother tongues, class, caste, religion, etc. In addition, as pointed out by Shermis and Di Vesta (2011), a teacher needs to keep track of student-learning through CAs which may happen "(a) *before* learning, (b) *during* learning, and (c) *after* learning" (p. 6).

2.3.3 Mutual-beneficialness

According to Shermis and Di Vesta (2011), "CA, at its best, is a positive collaborative activity between teacher and students" (p. 16). Earlier researchers (Angelo & Cross, 1993; Gipps, 1994; Dixon-Krauss, 1996; Shepard, 2000; Tierney, 2006, etc.) also share similar views. As mentioned earlier, this kind of assessment gives more space and opportunities to students to participate and contribute to the process of learning thereby providing them with a much-required sense of achievement. Students become more responsible and try to take charge of their learning. Generally, a teacher would like to achieve this. In addition, they get feedback on their teaching and students' progress. So what is evident here is an increased level of "collaboration and communication" (Steadman, 1998, p. 27) happening due to CA.

2.3.4 Formativeness

Marzano (2006), after reviewing the existing research on CA, says: "CA should be formative in nature." (p. 3). Though FA does not have any fixed definition (Black & William, 1998), Heritage (2010) manages to capture the distinctive features of FA in her description. According to her, it is "a process that occurs during teaching and learning and involves both teachers and students in gathering information so they can take steps to keep learning moving forward to meet the learning goals" (p. 8). To put it in other words, CAs often employ multiple measures like observation, self- and peer-assessments, journal, portfolio, surprise tests, quizzes, etc. to collect information about students' learning. So students do not have to worry about one final examination/test in which they have to perform well. Furthermore, they get an array of opportunities spread across an academic year to perform.

FA is discussed elaborately later in this chapter.

2.3.5 Context-specificity

Just as every learner is unique in some sense, a particular classroom may require a specific kind of assessment, different in some sense from one that is required by another. CA should be context-specific (Stiggins, Arter, Chappuis&Chappuis, 2004; Marzano, 2006; Dann, 2012). Since the stakeholders in the process include not only teachers and students, but also parents, curriculum and syllabus designers, textbook writers, policy-makers, etc., the demands of one such group are expected to vary from other such groups. Thus, it important to have CAs that cater to the demands of the stakeholders in that particular context. However, context-specificity cannot stand as an isolated characteristic different from the ones mentioned in the previous paragraphs. A student-centred, formative and mutually-beneficial assessment can be expected to be context-specific.

2.3.6 On-goingness

CA is a process that is always evolving. It happens simultaneously with classroom instruction (Moody &Stricker, 2009) and thus, is natural (Smith, Smith &Lisi, 2001). Continuous Comprehensive Evaluation (CCE) suggested by NCERT (2006) reflects this particular aspect of CA. Marzano (2006) cites findings of research projects undertaken by Fuchs and Fuchs (1986) and Bangert-Drowns, Kulik, Kulik and Morgan (1991) to claim

that higher frequency in formative CAs gives rise to higher student achievement (pp. 9 – 10). Russell and Airasian (2012) comment about the ongoingness of CA in no uncertain terms: "Teachers must continually observe, monitor, and review student behaviour and performance in order to make informed decisions" (p. 5).

2.3.7 Authenticity

An assessment can be termed as authentic if it is meaningful and rewarding, demands full participation on the part of students and assesses abilities and skills beyond the immediate use of the assessment (Frey, Schmitt & Justin, 2012). Going by this definition, CAs need to be authentic. It is important for students to apply the knowledge and skills learnt in the classroom in meaningful and real life contexts and this is precisely what an authentic assessment offers (Wiggins, 1989; Newmann, 1991; Brooks & Brooks, 1993). However, Anderson's (2003) argument that authentic assessment should be replaced by "assessment of authentic learning" (p. 73) cannot be ignored. The implication is that teachers must try to maximize the use of authentic tasks for instructional purposes so that students can be assessed fairly in authentic contexts.

Together, all the above-mentioned features of CA make what Angelo and Cross (1993) call "good practice" (p. 4). What they mean by this has to do with the harmonious integration of assessment and teaching. And alternative assessment techniques are of utmost importance considering the way they hold teaching and assessment together. But before looking at alternative assessment, it may be appropriate to present a brief review of research on CA in English Language Education (ELE).

2.4 Assessment of the English Language in Classroom: A Brief Review of Research

CA in ELE began long back in the 1980s as a reaction to standardized testing (Yang, 2007). However, systematic research into teacher-based assessments began only in 1990s (Cumming, 2009). But even after twenty years, the meaning of CA in ELE context carries considerable disagreements and variations (Rea-Dickins, 2007) and moreover, research on CA in ELE is lagging far behind those in other subjects (Davison & Leung, 2009). Leung (2005) reviews the most well-known of all early arguments and finds that the term 'CA' is associated with "alternative assessment" (Huerta-Macias, 1995), "authentic assessment" (Garcia & Pearson, 1994), "educational assessment" (Gipps, 1994), "formative assessment" (Black & William, 1998), "informal assessment" (Rea-

Dickins, 2001), etc. These terms and the associated definitions indicate different aspects of the process of assessment. As appropriately stated by Davison & Leung (2009), they "…all tend to be used to signify a more teacher-mediated, context-based, classroom-embedded assessment practice" and are often described as alternatives to "…traditional externally set and assessed large scale formal examinations" (p. 395).

2.4.1 Alternative Assessment

More than anything else, CA can be approached as an alternative to large-scale standardized tests. In some cases like that of India, it has been recognized as an alternative to centralized tests implemented in schools. In most educational contexts, alternative approaches to assessment have been sought with a view to giving equal opportunities to all learners to learn and perform. According to Hamayan (1995), many positive changes can be infused into the educational system, and learning can be enhanced using alternative approaches to assessment. However, researchers like Huerta-Macias (2002) and Farhady (2003) have also mentioned the difficulties lying with the large-scale effective practice of alternative methods in the assessment of language in classrooms. But Farhady (2003) goes on to talk about the many advantages of alternative assessment.

Apart from being innovative and fresh, alternative approaches follow a whole language approach, promote student-involvement and learning in communicative contexts and provide qualitative information about learners' developing proficiency (Farhady, 2003). They are "non-intrusive to the classroom" (Huerta-Macias, 2002, p. 339) because they can be easily integrated into the instructional practices in the classroom. So, students get a chance to improve their performance even while being assessed.

Rea-Dickins (2007) describes a list of alternative assessment procedures suggested by Brown and Hudson (1998). The list includes checklists, journals, video-tapes, portfolios, and self- and peer-assessment. But there were other approaches to CA before these procedures became popular. Dolan (1978) while discussing the problems with CA of language mentions a few assessment procedures like group cloze, discussions, etc. used for reading skills assessment. Cohen (1980) tries to bridge the gap between discrete-point testing and newly-emerging integrative approaches. But his conviction that testing is a separate activity that follows teaching does not help him with bridging. Furthermore, he suggests use of quizzes, discrete-point tests, dictation and a few integrated-skills

approaches for classroom *tests* (not *assessments*) which are secondarily used for CAs these days. However, Cohen advocates collection of information about student learning using multiple procedures which stand relevant even today.

One of the most comprehensive and earliest attempts made at putting together information related to CA was made by Genesee and Upshur (1996). Their suggestions are more of a practical nature and based on recent research on language acquisition and abilities. They discuss a wide range of assessment procedures starting with observations and conferences to dialogue journals and tests (p. 70). These alternative assessment procedures have become part of CA and are being refined by many current researchers. The aim has been to embed assessment with teaching (Short, 1993; Leung, 2005; Rea-Dickins, 2007). Black (2009), Lantolf (2009) and Falsgraf (2009) also emphasize continuous interaction between instruction and assessment.

There are efforts in ELE to bridge the gap between the standardized test and CA. Stoynoff (2012) thinks "DIALANG is a good example of progress in the effort to align large-scale language assessments more closely with CA purposes and practices" (p. 526). This time-impromptu assessment provides feedback on the examinee's strengths and weaknesses in using language skills. Stoynoff (2012) has authoritatively stated that the hegemony of large-scale single score-yielding multiple-choice tests is over (p. 527). Stoynoff is very much aware that sociocultural theories have found their way into the field of language testing and emerged potentially in the form of dynamic assessment. Studies by Poehner (2009) and Lantolf (2009) indicate that dynamic assessment offers a framework for CA of language and it adds to the cause of CA.

2.4.2 Authentic Assessment

Another alternative procedure of assessment that has got some research attention and been used under the umbrella term CA is authentic assessment (O'Malley & Pierce, 2002; Fulcher & Davidson, 2007; Delli Carpini, 2009). According to O'Malley and Pierce (1996), it is based on the principles of constructivism and should be preceded by a similar approach to teaching. They opine that authentic assessments "should provide the students with opportunities to construct responses and apply their learning to problems that mirror their classroom activities in authentic ways" (p. 10).

2.4.3 Formative Assessment

Formative and diagnostic assessments also feature in CA. FA is an offshoot of the point of intersection between language acquisition and assessment (Theodoropulos, 2011). Formative aspects of assessment in classroom have been discussed and emphasized in different ways. Purpura (2004) calls it "learning-oriented language assessment" and Rea-Dickins (2008), "assessment as learning". There is no disagreement about the fact that FA "takes place during rather than at the end of a course or programme of instruction" (Association of Language Testers in Europe, 1998, p. 142) and that the assessment results can be used by the teacher to scaffold individual students with their learning. On the other hand, diagnostic assessments have to do with "discovering a learner's specific strengths or weaknesses" (ALTE, 1998, p. 142). Huhta (2010) thinks FA is more grounded in a prescribed course/textbook/curriculum than dynamic or diagnostic assessment which is often based on a theory/model/framework (p. 473).

2.4.3.1 Formative vs. Summative Assessment

The principles and practices of FA can be easily understood when they are contrasted with those of summative assessment (SA). Bloom, Hastings and Madaus (1971) distinguished between teacher-developed assessments for classroom purposes and assessment for judging students' progress by terming them 'formative' and 'summative' respectively. Ever since, SAs are employed to judge students' performance against a set of criteria. They are achievement-oriented and required to be valid and reliable as the judgments often result in scores and grades (Pachler & Redondo 2014). Often, students are not provided with any feedback about their progress after SAs. Even if they get some information about their performance, it is a little too late (Popham, 1999). Another argument that goes against SA is that it is not really based on what happens in the classroom (Shepard, 2001). The above mentioned problems are addressed by FAs as they provide learners with the necessary feedback about their progress and weaknesses (Lewy, 1990) and help both teachers and learners adjust their teaching and learning accordingly (Perie, Marion & Gong, 2007). So FA actually promotes effective learning. It is discussed below as 'assessment for learning'.

2.4.3.2 Assessment for Learning

One of the fundamental aims of FA is to utilize assessment to support learning, i. e. assessment for learning (AFL). According to Brookhart (2010), FA is an 'ongoing process' involving students and teachers aiming to:

1. Focus on learning goals.

2. Take stock of where current work is in relation to the goal.

3. Take action to move closer to the goal.

(Brookhart, 2010, p. 3)

What Brookhart talks about has little to do with assessment that is used for judgment purposes. The focus is entirely on learning. According to Black and William (2009), through FA "evidence about student achievement is elicited, interpreted, and used by teachers, learners, or their peers, to make decisions about the next steps in instruction" (p. 9). The collected evidence forms the core of AFL. According to the Assessment Reforms Group (2002), AFL "is the process of seeking and interpreting evidence for use by learners and their teachers to decide where the learners are in their learning, where they need to go and how best to get there". Though the collection of evidence suggested in this definition leads to decision-making, the decisions are directed towards improving classroom teaching and learning. The list of activities included under AFL category and often suggested by researchers and assessment experts for carrying out FAs in the classroom include encouraging thinking, being sensitive and supportive, motivating, encouraging a variety of answers, employing peer- and self-assessment, offering scaffold and suitable feedback, etc. (Florez & Sammons, 2013).

2.4.3.3 Integrated Assessment

The above discussion suggests that FAs should be embedded in classroom teaching and interaction (Pryor & Crossouard, 2008). The integration of teaching and assessment can yield rich dividends in terms of student achievement. The findings of study by Wiliam, Lee, Harrison and Black (2004) empirically support this claim about the integration. Integration of instruction and assessment can lessen the burden of assessment on both students as well as teachers. While discussing FA, Shavelson (2006) talks about three ways in which teachers can assess while teaching. They are: *on-the-fly*, *planned-for-interaction*,

and *formal and embedded in curriculum*. The first kind is one in which a teacher exploits a sudden opportunity arising in a classroom situation for assessment purposes. The second one involves a somehow deliberate attempt and prior preparation to utilize situations in which students' suddenly face some knowledge-gap. The third kind, as the name suggests, is a scheduled attempt to practice assessment in the classroom at regular intervals. A teacher requires some amount knowledge and skill about FA to practice any of the above mentioned kinds of FA in their classroom.

2.4.3.4 Tools for Formative Assessment

FAs make it possible for the teacher to 'individualise the support' (Mathew, 2012) for students so that they can learn the targeted skills effectively. Thus, a variety of strategies and tools are used in FAs. Brookhart (2010, p. 12) lists five general strategies that guide FAs:

- reflection questions
- indicator systems
- logs or diaries
- review of one's own work against criteria
- goal setting or action planning

These strategies overlap with some of the alternative methods of assessment, which are often associated with FA, suggested by Genesee and Upshur (1996). They include the following:

- observation in the classroom
- portfolios
- conferences
- journals
- questionnaire
- interviews

The only difference between the above mentioned two lists is that the former is a set of student tools and the latter, teacher tools. A combination of both can yield desired results.

2.5 Maintaining the Quality of Classroom Assessment

Even though alternative assessment procedures dominate CA, there has been emphasis on standard-based CAs (Leung & Teasdale, 1997b; Breen *et al.*, 1997; Brindley, 1998, 2001; Arkoudis&O'Loughlin, 2004; Leung, 2007, Llosa, 2007, 2011). McKay (2006) and Llosa (2011) report that many countries have adopted English language proficiency (ELP) standards. But, Llosa (2011) also points out that using the result of standard-based ELP tests for formative purposes is a challenge for teachers. The implication is that levels of language proficiency should be very clearly defined in a context-specific manner and teachers should be made aware of what the 'mastery' of each level means.

If maintaining standards is a problem in CAs, the quality of CA in ELE has also been a matter of debate. Stoynoff (2012) cites researchers who contend that CA should follow the principles of language assessment as it is done in case of large-scale standardized tests (Cohen, 1994; Huerta-Macias, 1995; Brown & Hudson, 1998, etc.). He also cites the view of another group (McKay, 2006; Rea-Dickins, 2008 and Davison & Leung, 2009) which thinks that it is not possible to follow the aforesaid principles for CA.

As standardized tests are going out of favour, there is a great deal of support for what Bachman (2007) calls "performance-based language assessment" (p. 55) and "interactional approach to language assessment" (p. 57). In such cases, the validity, reliability and other qualities can be evaluated on the grounds of authentic information about learners' progress collected through multiple methods. The array of assessment methods used in CA adds to the objectivity of information collected about learners' language ability. Having said this, it may be difficult to ignore concerns like construct and content validity (Messick, 1994), inconsistency in rater judgements (Mehrens, 1992), very little scope for generalizing performances (Dunbar, Koretz& Hoover, 1991), etc. related to such assessments. The origin of these arguments can be traced to research on the impact of external tests and reporting on classroom practices (Hill & McNamara, 2012). Though there are very few studies on the actual process of CA (Hill & McNamara, 2012), yet the arguments regarding the utility of CA cannot be discounted. In instructional contexts where learning is held beyond everything else and where individual learners matter, CA is a potential help. It can generate positive wash back which, in turn, can help in facilitating learning (Stoynoff, 2009). However, as pointed out by Rea-Dickins (2007) there is little information about what goes into the making of a quality language assessment. Acknowledging that 'quality' is a sociocultural variable, the present study

makes an effort to define 'quality' in the context of the society, curricular policies, schools, classrooms, teachers and students under scrutiny.

2.6 Teacher Role in Classroom Assessment

The success of CA depends largely on the teacher. According to Cumming (2009), teachers need "to become familiar with, use, and further develop the broad range of resources, principles, practices, and research findings that have recently emerged about CA" (p. 1) to be effective practitioners. They have been called "agents" of assessment (Rea-Dickins, 2004), responsible for observing, evaluating and interpreting learners' progress. Brown (2004) thinks they are more like "tennis coaches" monitoring student learning. So the teacher is continuously involved in the process. However, teacher roles in CA vary with the form and procedure of CA (Cheng & Wang, 2007). Yin's (2010) study suggests that teacher's thinking also has an impact on their CAs. He talks about "assessment cognitions" which comprise "strategic cognitions" and "interactive cognitions", to suggest that teachers' assessment practices are determined to a great extent by these factors.

A teacher's involvement in the process is perhaps the single most important factor in CA (Leung, 2005). It is a factor that enables a teacher to grow professionally. Whether the assessment is called teacher-based or classroom-based, performance-based, formative or dynamic, the involvement of the teacher is always emphasized. Here the involvement can be with students as in case of dynamic assessment (Poehner, 2009). Teaching and assessment happen simultaneously and knowledge is co-constructed. In case of student-self assessment, teachers play the role of a facilitator (Bullock, 2011). Yet another role is demanded of a teacher in case of peer-assessment. But not many attempts have been made to specify and define teacher and learner roles in specific assessment CA situations. Mok (2011) thinks the roles of teachers and students in peer-assessment should be re-defined.

In ESL contexts like that in India where performance in summative assessments is still considered extremely important and where mastering English language skills is crucial to students' future, teachers need to have what Edelenbos and Kubanek-German (2004) call "diagnostic competence". It can be defined as "the ability to interpret foreign language growth in individual children" (p. 259). But this perspective is a little narrow (Seong,

2011) and thus, should be expanded so that it can accommodate factors like using the interpretation of assessment results for propelling learning.

As the teacher is expected to perform multiple roles, all of which are challenging, they should be given professional training and other kinds of help so that they can perform their assessment duties effectively. The assessment scenario, curricular and organizational demands, learners' language needs, their sociocultural background, etc. should be analysed thoroughly to arrive at the training needs of teachers. Such a framework is developed for the current study later in the chapter.

2.7 Assessment Literacy: Definitions

The term "assessment literacy" was perhaps used by Stiggins (1991) for the first time to refer to the knowledge and skills required by teachers, syllabus designers, textbook writers, test designers and policy makers to define, design, interpret and use tests for various curricular purposes. However, one of the earliest efforts to list out a set of teacher competencies in testing happens to be a joint venture by the American Federation of Teachers, the National Council on Measurement in Education and the National Education Association (1990). These were meant to be used both in in-service and pre-service teacher education programmes and included the following:

- choosing appropriate assessment strategies
- developing appropriate assessment methods
- administering, scoring, and interpreting assessment results
- using assessment results for making decisions about individual students
- developing valid grading procedures
- communicating assessment results to stakeholders
- using assessments in a democratic and fair way

These standards of teacher competence are, however, found outdated by Brookhart (2011). She thinks that these standards are not congruent with today's "conceptions of formative assessment knowledge and skills" and "the knowledge and skills required to successfully work" in the current context (p. 3). The implication is that the concept of assessment literacy is a dynamic one. In case of performance-based CAs, a teacher has to interpret students' performances qualitatively in addition to analyzing their test scores (Kane, 2006). Brown (2008) tries to come up with a broader definition. According to

him, "(a)n assessment-literate teacher is one who creates, chooses, administers, interprets, responds to, records, and reports assessment information in such a way that those decisions can be shown to be adequate and appropriate" (p. 286). However, Brown does not indicate the specificities in relation to contextual demands in his definition. The concepts "adequate" and "appropriate" are context-specific.

White (2009) defines assessment literacy as "the kinds of assessment know-how and understanding that teachers need to assess *their* students effectively" (PPT on www.slideshare.net). Though in this definition, the word "their" covers the context issue, the wider impact of assessment and teacher beliefs is still not covered in the definition. According to Gamire and Pearson (2006) (as referred to in Torrie& Van Buren, 2008, p. 41), the ability to assess includes "knowledge about assessment, critical thinking and reflective judgement skills, and capabilities in the use of content knowledge to solve practical problems". This definition adds a valuable dimension to the concept. Curtz (2007) further adds that one requires to have the ability to self-assess and self-reflect to assess others. All the above-mentioned abilities, knowledge-bases, skills, etc. comprise a solid understanding of what go into making sound assessments (Stiggins, 2002; Popham, 2004; Volante& Fazio, 2007). Popham (2009) suggests a series of content areas that comprehensively cover most of the concerns shown by researchers in the area.

2.7.1 Assessment Literacy in English Language Education

The above discussion on the definitions of assessment literacy is more related to educational assessment than to assessment in ELE. To find a comprehensive definition in the field of ELE is a tough challenge . Quite strangely, this term which is very relevant to current practices in language testing does not feature in dictionaries like *Dictionary of Language Testing* (1999) and *Encyclopedic Dictionary of Language Testing* (2002). This absence indicates that assessment literacy is yet to get fully recognized by the language assessment research community. So the aim here is to construct a comprehensive definition with the help of existing ones. It must also be made clear that since the current study focuses on English teachers, the discussion on language assessment literacy will concentrate on, if not be confined to, the requirement of teachers.

Brindley (1997) was one of the earliest to use the term "assessment literacy" in the language education context though language testing courses and their components were discussed earlier by Bailey and Brown (1996). According to Brindley, teachers need to

have "a wide range of skills" to formally assess their students. This set of skills is quite similar to the ones prescribed as Standards for Teacher Competence in Educational Assessment of Students. But a few years later, Brindley (2001), while studying teacher-training courses, added a few more components to his list. These components include social context of assessment, definitions of language proficiency, construction and evaluation of assessments and knowledge about the interaction between curriculum and assessment. But there has been uncertainty about what exactly teachers and other stakeholders should have to become assessment literate. In this connection, Malone (2007) says, "[t]here is no consensus on what is required or even needed for language instructors to reliably and validly develop, select, administer and interpret tests" (p. 225).

With the rise of sociocultural theories and consequently, constructivist approaches in ELE, language assessment has also taken a "social turn" (McNamara, 2006) and been under the influence of the same theories (Inbar-Lourie, 2008). According to Inbar-Lorie, who makes a thorough review of how conceptions of learning and assessment have been redefined in the last decade, assessment literacy may refer to "[t]he capacity to ask and answer critical questions about the purpose for assessment, about the fitness of the tool being used, about testing conditions, and about what is going to happen on the basis of the results." (p. 389). This definition tries to accommodate most of the abilities and covers the "why", "what" and "how" (Inbar-Lorie, 2008, p. 390) of assessment. But it demands a lot of further clarifications regarding the terms used. The "why" may have to do with societal, educational and political needs (Brindley, 2001), whereas the "what" can be related to the understanding of theories and cultural perceptions of language, pedagogical content knowledge and curricular mandates; and the "how" can be associated with approaches, methods and techniques of assessment.

Even the broad definitions of language assessment literacy (LAL) do not offer much concreteness, flexibility and inclusiveness. The research shows it is better left floating. Davies (2008), after an extensive review of teaching of language testing, comes to the conclusion that any definition of assessment literacy should be applicable to the corresponding period and educational context. According to Davies (2008), assessment literacy for language educators should include "skills", "knowledge" and "principles" of language testing. Here, skills encompass "necessary and appropriate methodology"; knowledge "offers relevant background in measurement and language description, as well as in context setting"; and principles "concern the proper use of language tests" (p. 328).

McNamara and Roever (2006), Spolsky (2008), Davies (2008) and Taylor (2009) prefer a broad view of LAL which considers life, society and education. Thus, LAL can be placed on a plane that comprise "(a)n appropriate balance of technical know-how, practical skills, theoretical knowledge, and understanding of principles, but all firmly contextualized within a sound understanding of the role and function of assessment within education and society" (Taylor, 2009, p. 27).

Malone (2011) looks at LAL from another angle and believes that as assessment goes hand-in-hand with learning, AL should entail "(a)n understanding of measurement basics related directly to classroom learning" (p. 1). This definition specifically focuses on assessment in the classroom. Fulcher (2012) shows reservations against any narrow definition of LAL and tries to build on Davies' concept of LAL. He warns against the preference for "individual or group perceptions" at the cost of "validity (or ethical) arguments" (p. 117) and comes up with "an expanded definition" (p. 125) that almost embodies all the existing definitions of LAL and includes broad categories like "contexts, principles and practices" (p. 126).

2.8 An Eclectic Framework of CLAL for the Current Study

After surveying most of the existing definitions of LAL and related lists of abilities, skills, etc., certain things have been found to be common in all of them. Coombe, Troudi and Al-Hamly (2012) very aptly summarise, "[t]hose who are assessment literate understand what assessment methods to use in order to gather dependable information about student achievement, communicate assessment results effectively, and understand how to use assessment to maximize student motivation and learning" (p. 25). Keeping these commonalities in mind, an eclectic framework of LAL has been developed for this study. The definition of LAL generated through this framework may appropriately represent the required LAL for secondary school English teachers teaching English language in different parts of India.

Before arriving at the actual framework, certain facts related to CA of the English language in secondary schools need to be made very clear:

- In the absence of any separate policy for ELE in schools, assessment of English language in schools, especially at secondary level, is determined by a common educational assessment policy.

- Most of the secondary school English teachers do not have Bachelor's or Master's degree in English. Even otherwise, syllabuses of B. A./ M. A. (English) courses across the country are heavily literature oriented.

- The teacher training institutes lack adequately qualified and trained Language Teacher Trainers.

- There is no special policy in the country for English teacher preparation.

- The value of year-end summative tests often overrides parental and institutional requirements.

- There is virtually no official data available about the number of English teachers practising in schools, their educational qualification, training, experience, etc.

- If an English teacher wants to grow professionally, he/she needs a very high level of intrinsic motivation.

- Quality research of international standards on Second Language Teacher Education is rarely undertaken and published by researchers in India.

- Learners need command of English language skills to get good jobs in future.

- But the English language skills of school teachers have been under the scanner for long (NCFTE, 2010).

- CA as a policy came to exist in the form of CCE but lack of teacher-preparedness affects it badly.

- CA in large and crowded classrooms can be really taxing if not handled with some kind of expertise.

The following model takes into account the aforesaid assessment scenario in India. Teachers' CLAL is shown to be inseparably related to the official needs, learners' language needs and teachers' own professional needs:

```
                                    ┌─────────────────────┐
                          ┌──────────│ ● Understanding policy-
                          │          │   related and
              ┌───────────────┐      │   organizational demands
              │ Official needs │─────│ ● Acting on them
              └───────────────┘      └─────────────────────┘
                      ⇕
┌──────────────┐  ┌────────────────┐  ┌─────────────────────┐
│ Teachers' CLAL│──│ Learners' language│──│ ● Awareness about
└──────────────┘  │    needs        │  │   learners' language
                  └────────────────┘  │   requirements
                      ⇕               │ ● Meeting these
                                      │   requirements
                  ┌────────────────┐  └─────────────────────┘
                  │ Teachers' professional│ ┌─────────────────────┐
                  │    needs        │──│ ● Awareness of
                  └────────────────┘  │   classroom teaching and
                                      │   assessment abilities
                                      │ ● Equipping self with
                                      │   the same
                                      │ ● Motivation to learn
                                      │   and positive attitude
                                      │   towards professional
                                      │   growth
                                      │ ● Appropriate beliefs
                                      │   about assessment
                                      └─────────────────────┘
```

Figure 2.1: An eclectic framework showing the CLAL of secondary school English teachers in India

The practicing English teachers in secondary schools need to be assessment literates in a local way, rather than in a pan-Indian way. The gap among different types of schools (CBSE, ICSE and State Board) combined with their financial situation, to a large extent, determines the quality of teachers. So it will be unfair to compare secondary school English teachers in a backward state like Bihar with those in Delhi or Gujrat. Thus, classroom language assessment literacy of English teachers in secondary schools can be defined as the ability to understand the meaning of language proficiency, the utility and role of language assessment, especially CA in the English classroom, the societal, organizational and individual student needs in terms of the English language, and plan, construct, use, interpret and evaluate classroom assessment of the English language so that their students can learn to be independent and effective users of the language. It also includes possessing appropriate beliefs about assessment.

The above working definition of assessment literacy keeps certain components flexible so that it can be used in contexts similar to the one under investigation. This definition will be used in the last part of the next chapter while presenting the theoretical framework for the current study.

2.9 Conclusion

After examining different aspects of CA and LAL, it is found that they both need to be used flexibly by stakeholders so that the language needs of learners can be catered to by using the available resources in their own context. In the next chapter, research on assessment training in SL language teacher education programmes, professional development of SL teachers in assessment, impact of in-service programmes on teacher knowledge and beliefs, teacher change, etc. is discussed with reference to the particular questions addressed in this study.

CHAPTER 3

TEACHER DEVELOPMENT AND CHANGE WITH SPECIAL REFERENCE TO LANGUAGE ASSESSMENT

3. 1 Introduction

In the previous chapter, the discussion centred around CA and LAL. This chapter examines the research literature related to SL in-service teacher education and development and teacher change with special reference to language assessment. The chapter has two parts. The first one focuses mainly on in-service language teacher development (TD) and teacher change, whereas the second part focuses on relevant research on in-service language TD and teacher change in relation to assessment. The first part includes review of research related to in-service TD, use of terms in the field, teacher learning, teacher belief, guiding principles for developing effective TD programmes, short TD courses and teacher change. The second part comprises review of research literature focusing on assessment in teacher education programmes, TD in assessment, teacher beliefs about assessment and changing assessment practices through TD programmes. The two sections provide a solid theoretical foundation for the study, especially for the designing of the intervention study and the analysis and interpretation of the research data.

3.2 In-service Education and Development of Language Teachers

Teacher learning is a continuous process. Changing societal demands, developments and innovations in the field of language education and improvements in educational infrastructure across the globe are some of the main reasons why language teachers need to keep updating themselves professionally as long as they are in the profession. Teachers can grow professionally by voluntarily participating in professional development (PD) activities and being a part of PD programmes provided by the school management and the government. Whether it is top-down (as in case of activities prescribed by the school) or bottom-up (as in case of voluntarily undertaken ones), any such activity aims to bring about constructive changes in teachers' pedagogical skills, knowledge, thinking, beliefs and performance to facilitate language learning among their learners.

Though it is necessary to follow pre-service programmes with in-service programmes (Cimer, Cakir & Cimer, 2010, p. 31), it is difficult to find many follow-up in-service programmes. Moreover, for no apparent reason, the research on in-service education of language teachers is not as intense and widespread as that on pre-service teachers (Mann, 2005; Wright, 2010; Vefalı & Tuncergil, 2012).

While learner factors get highest priority in any education programme, several other factors need to be considered before finalizing the components of an in-service programme. The first of these can be teachers' individual needs and available opportunities in the given social set-up (Ballantyne, Sanderman, & Levy, 2008). The second is creating space for individual and collaborative learning as they benefit the teacher in gaining expertise in the profession (Desimone, Porter, Garet, Yoon, & Birman, 2002). The next important factor concerns the duration and frequency of in-service training (Fullan, 2001; Mata, 2012). The last major concern is the integration of teachers' personal experiences into the programme to make the programme more relevant to teachers (Erickson & Anderson, 1997). Moreover, there should be a strong connection between pre-service and in-service training programmes so that TD becomes a continuous process. In addition, certain areas like assessment that do not get adequate attention during pre-service training can and should be covered in the course of in-service training programmes.

3.2.1 Terms and their Meanings in In-service Teacher Education

It is necessary to look at how different researchers have defined terms related to courses aiming to promote in-service professional learning of teachers as this dissertation deals with an in-service short course for enhancing teachers' CLAL. According to Craft (2000), terms like Continuous Professional Development (CPD), Professional Development (PD), In-service Education of Teachers (INSET), TD (TD), etc. are often used to mean the teacher learning that is triggered through a variety of in-service training programmes. Researchers are found to be divided in the way they prefer to use terms to refer to different kinds of in-service teacher professional support programmes. It is necessary to review some of these views as this dissertation deals with a short term intervention programme aiming to enhance teachers' CLAL.

Ur (1997) and Ohata (2007) have cited the three models of teacher learning: *applied science*, *craft* and *reflective*, advocated by Wallace (1991) to differentiate between 'training' and

'development'. Ur and Ohata point out that while the characteristics of the first two models match the features of 'training', *reflective* model encompasses the principles of 'development'. Ur (1997) reviews the way the two terms— 'training' and 'development'— have been used by researchers like Bolitho (1986), Edge (1986), McGrath(1986), Freeman (1990), etc. and lists a number of differences between the meanings conveyed through the use of terms. According to that list, 'training' has to do with a top down and predetermined programme developed by experts. It is often seen as a stand-alone, disempowering and skills-focused model. In contrast, 'development' refers to a teacher's voluntary attempt to grow professionally through an evolving, need-based, teacher-driven, holistic and empowering model of development that involves the teacher as a whole. Ur (1997) finds both categories incomplete and unsatisfactory and suggests an integration of components from both the categories. A similar view is shared by Edge (2003) who, though he found differences among the use of terms like 'education', 'training' and 'development', clearly states that it is not possible to create water tight compartments for the terms and advocates bringing together the useful principles conveyed by these terms while helping teachers to grow as professionals.

Mann (2005), in his state-of-the-art article on the language teacher's development, very distinctly defines the use and meaning of the aforementioned terms and claims that the *training* and *education* models of development are shared by Americans and the *development* view is subscribed to by British researchers.

The above review suggests that there is no strict terminology as to what an in-service programme should be called as long as it helps a teacher grow professionally. In this connection, Richards (2008) quite appropriately states that "the contrast between training and development has been replaced by a consideration of the nature of teacher learning, which is viewed as a form of socialization into the professional thinking and practices of a community of practice" (p. 159).

As the current study is carried out in India and involves Indian school teachers, a review of how these terms have been used in the Indian context is presented in the following sub-section.

3.2.1.1 A Brief Review of Indian Perspective

In most of the Government documents, 'in-service training', 'in-service education' and 'INSET' are used to refer to short-term and long-term programmes meant for promoting

teacher learning and professional growth. In *National Curriculum Framework for Teacher Education* (2009), the term 'TD' has been used only once and 'INSET' has never been used. However, 'INSET' has been used in an NCTE document— *Study of Impact of In-service Teacher Training under SSA on Classroom Transaction* (2012)— to refer to in-service training and education of teachers. What is evident here is that 'TD' is not used for top-down and government sponsored programmes in India.

Sen (2002) in her Ph. D. dissertation discusses models of teacher education and discusses research related to the same. She stresses on 'participatory teacher training' and suggests that "training is based on reflection and introspection, and the teacher's needs, problems, status, roles, etc. are not presented by outside observers or experts" (p. 17). She uses terms like 'professional development' and 'continuous professional development' along with 'in-service training/ education' alternatively.

Barrett (2010) uses the term 'INSET' to refer to in-service teacher education programmes conducted in India. In her discussion, however, she focuses on many of the components like teacher identity, needs, beliefs, lives, teaching context, etc. that are often discussed under 'TD' by many British researchers.

Mathew (2014) points out concrete differences among the terms 'teacher training' (TT), 'teacher education' (TE) and 'TD'. According to her, *TT* refers to "familiarising student teachers with techniques and skills to apply in the classroom", *TE* involves "teachers in developing theories of teaching, understanding the nature of teacher decision making and strategies for self-awareness and self-evaluation..." and *TD* has to do with "a voluntary process, on-going, bottom-up since the starting point is the teachers' own experience where new information is sought, shared, reflected on, tried out, processed in terms of personal experience and finally 'owned' by the teachers". Here, she is referring to 'training' as something to do with pre-service teachers and 'education' and 'development', with in-service teachers.

Padwad and Dixit have been involved in research on teacher education in India. In articles by Padwad (2011), Dixit (2011) and Padwad and Dixit (2014), they make use of the term Continuous Professional Development (CPD) to refer to in-service programmes outside those that are offered by government. Much like Mathew, they seem to prefer 'development' over 'training'.

In the current study, the intervention programme was not part of any formal in-service teacher education programme. It took into account the CA needs of the participating teachers. Though the teachers were encouraged to participate in the programme by their school management, they had the freedom of dropping out of it. Moreover, the teachers had a say in how the programme was conducted and they did not have to abide by the trainer-trainee relationship with the instructor. Thus, the intervention programme, which quite evidently drew on features ascribed to both 'training' and 'development', was called a TD programme, not INSET. This is a reason why reviews of both TD and other in-service programmes are included in the section on the impact of such programmes on teacher learning.

3.3 In-service Learning and its Impact on Teachers

Ideally, any in-service teacher professional development programme should be a part of a continuous professional development (CPD)-continuum. It should be grounded in the pedagogic context in which the target group of teachers practise, and take into account their prior experiences of teaching, learning and professional training and education. The complexities involved in constructing such programmes are evident from the low rate of success of these programmes as reported by Veenman, Tulder & Voeten (1994), Pacek (1996), Waters (2006), and a few others. These researchers have also pointed out the difficulties in assessing the impact of these programmes. The impact of any such programme can be on any one of or all the areas like teacher competence, knowledge, belief, attitude and practice. So examining the impact on each area poses an enormous challenge to any researcher.

Teacher learning through in-service programmes can be interpreted through sociocultural theories of learning proposed by Vygotsky and Bruner. According to Vygotsky (1978), learning precedes development, and learning is a social phenomenon and mediated by sociocultural artefacts. He also claims that learning happens through social interactions and is propelled by 'scaffolding' (Bruner, 1975, 1978). If these theories are applied to teacher learning through such programmes, it can be claimed that the success of such a programme depends to a large extent on how thoroughly it involves the target group of teachers during its construction stage, how well it takes into account sociocultural factors like their background, beliefs, experiences, etc., how convincingly the teacher educator provides necessary support to teachers in blending the new knowledge and skills with their existing knowledge and skills, and how much teachers

find the new inputs useful and applicable in their respective contexts. In short, a successful in-service teacher education programme is one that can help teachers 'internalize' (Langford, 2005) the input provided during the programme.

The number of studies on in-service teacher education is much less than those on pre-service education. Moreover, in-service studies often focus on various aspects of teaching—such as beliefs, knowledge, skills, etc. Sen (2000) while working with and on in-service teachers found that teachers do go beyond in-service programmes and think about the problems they face and search for solutions. She also asserts that a PD programme that addresses the trainees' problems and is grounded in their context has more chance of being effective. Lamie (2002) studies the impact of an in-service teacher education programme on a group of Japanese English teachers and reports that there were improvements in their "attitudes, espoused methodology and classroom practice" (p.135). She highlights how 'personal attributes', school culture, 'feedback', the existing awareness of teachers, the quality of training, etc. shaped the impact of the programme. Padwad and Dixit (2008) try to assess the impact of teachers' participation in professional learning communities on their thinking and attitudes towards classroom problems. They report that teachers find the participation quite useful and it makes them think positively about their classroom problems and search for solutions to these problems. Singh (2011), on the basis of a study conducted in India, finds that context-specific and indigenous models of in-service professional development programmes can be effective and induce confidence among teachers. Borg (2011) tries to trace the impact of an 8-week intensive in-service programme on teachers' beliefs. His findings indicate positive changes in the beliefs of the participants. However, he quite rightly states that the 'impact' of any in-service programme depends on how the term 'impact' is operationalized. Cheung (2013) also reports positive impact of a workshop on teachers' teaching skills and attitudes. He cites Opfer, Pedder and Lavicza (2011) and Avalos (2011) and points out that it may not be possible for teachers to make use of their professional learning immediately in their own classroom and school as 'school culture' (Mathew, 2006) and institutional management remain two of the important deciding factors.

Studies on the impact of government sponsored relatively large-scale in-service programmes in India (Subramanian, 2001; Raina, 2005; Eshwaran & Singh, 2008; Yadav, 2012; Mohanty, 2014; etc.) are not being included here since the current study is not

directly related to such programmes and their impact. The current study involves a small-scale TD programme that tried to cater to the needs of a group of only six teachers.

Since TD is a continuous process and the stage of internalization takes time to be reached, focus should be more on long-term impact of such programmes (Mathew, 2006; Luchini, 2010). However, as already pointed out in one of the above paragraphs, it may not always be possible to trace the long term impacts. Moreover, as pointed out by Gebhard (2006), "small changes can have big consequences" (p. 23). The small changes found in teachers' beliefs, practices, knowledge, etc. cannot be ignored. They can be potential clues about long-term impacts. Harbon (2007), Chinda (2009), Cook (2009), Harvey, Roskvist, Corder and Stacey (2011), Giraldo (2014) and Wang (2014) report positive change in attitudes towards their job and capacity to perform well as a result of a short-term programme. However, all the above mentioned researchers interpret the positive results with caution. This indicates that such positive impacts must be interpreted as a part of teachers' continual development process. Overgeneralization of the gains obtained through such programmes can be misleading and counterproductive, and thus, should be avoided.

Before arriving at the features of an effective in-service TD programme and laying out its principles for the intervention study carried out for the current research project, it is relevant to look into the research on teacher knowledge and belief, which is closely related to teacher learning.

3.4 Teacher Knowledge

In the last fifty years, there have been many changes in the field of Second Language Teacher Education (SLTE). These changes have been prompted and fed by research on linguistics and language teaching methodology and have led to the widening of knowledge-base for teachers (Pawlak, 2011). Earlier, from the positivist epistemological perspective, teacher knowledge was considered to be external to the teacher and could be transferred to the teacher through reading of theories, lectures by experts in the field and other external forces (Johnson, 2009). This perspective was based on the assumption that teaching should follow a set behavioural pattern that could result in student achievement in the form of high scores in tests. The claims to scientific status; the overemphasis on product-oriented approach to teaching and learning; and the view of knowledge as something that is objective and transferable from abstract to concrete real life situations

were challenged and criticized by socioculturalists (Carr and Kemmis, 1986; Johnson, 2009). They pointed out that crucial factors like the sociocultural context of teaching and learning, the identity of the teacher as an individual, and the complex process of teaching and learning were not considered by the positivists.

Subsequent attempts were made to redefine teacher knowledge taking into account factors ignored by the positivists. The most well-known of such attempts was made by Shulman (1986) whose definition of teacher knowledge-base was much more holistic and included the following categories:

- content knowledge
- general pedagogical knowledge
- curriculum knowledge
- pedagogical content knowledge
- knowledge of learners and their characteristics
- knowledge of educational contexts
- knowledge of broader educational aims

(Adapted from Shulman, 1986, p. 8)

With the emergence of research on teacher cognition, which is defined by Borg (2003) as "the unobservable cognitive dimension of teaching – what teachers know, believe, and think" (p. 81), in the last two decades, teacher knowledge has been re-conceptualized, and now it includes components like teacher personality, their existing knowledge and beliefs, their developing concept of teaching and the way teachers learn to adjust to different institutional contexts. According to Johnson and Golombek (2011), who affirm their faith in the social nature of teaching and learning, "knowledge for teaching must be understood holistically, and the interdependence between what is taught and how it is taught becomes crucial to both the processes of learning-to-teach as well as the development of teaching expertise." (p. 3). However, as pointed out by Freeman (1994), very little research has been done on the nature of knowledge required by the teacher to do what they do in the classroom.

3.4.1 Teacher Knowledge in Teaching Practice

Research on how teacher knowledge gets reflected in teaching practice has been inconclusive though the area has drawn a lot of research interest in recent times. The

inconsistent and complex relationship between what Cochran-Smith and Lytle (1999) call "know more" and "teach better" could be a reason for the inconclusiveness. In other words, there is no uniform relationship between knowledge about teaching and the practice of teaching. However, the impact of certain components of teacher knowledge like pedagogical content knowledge (Ball, Thames & Phelps, 2008), knowledge of educational context (Burns, 1996; Crookes & Arakaki, 1999), knowledge of learners' attitude and perceptions (Savignon & Wang, 2003), and knowledge of curriculum (Cumming, 1989) on their professional practice has been explored in applied linguistics research. But as noticed by Mewborn (2001), individual differences among teachers still remain a powerful variable that forbids any kind of large scale generalization in relation to impact of teacher knowledge on practice. The sociocultural approaches to teacher education, as believed by Johnson (2009), emphasize acquisition of new knowledge by the teacher as a dynamic and continuous process in which teaching and learning act as mediators. This social constructivist view of knowledge looks at the development of knowledge-base as a process of construction. This process is expected to vary from one teacher to another and one classroom setting to another.

3.5 Teacher Beliefs

Like teacher knowledge, teacher beliefs shape teachers' instructional decisions and practice. Due to the individual and personal nature of beliefs, it is almost impossible to find one single universally accepted definition of teacher beliefs. One of the most frequently quoted definitions is proposed by Borg (2001) who defines belief as "a proposition which may be consciously or unconsciously held, is evaluative in that it is accepted as true by the individual, and is therefore imbued with emotive commitment; further, it serves as a guide to thought and behaviour" (p. 186). A more condensed definition is offered by Richardson (1996). According to him, beliefs are "psychologically held understandings, premises, or propositions about the world that are felt to be true" (p. 103). Attitudes and beliefs are often used alternatively (Pajares, 1992; Richards, 1996; Liu, Wang, Nam, Bhattacharya, Karahan, Varma & Roehrig 2012), but Allport (1967) defines attitudes as "a mental and neural state of readiness, organised through experience, exerting directive or dynamic influence upon the individual's response to all objects and situations with which it is related" (p. 8). However, in this study, both beliefs and attitudes are used to mean the same set of constructs, and this part of the review concentrates only on language teacher beliefs in relation to their pedagogical practices.

Research on second language teacher beliefs has been on the rise in recent times (Woods, 1996; Borg, 2003, 2006, 2011; Andrews, 2003; Farrell & Particia, 2005; Theriot & Tice, 2009; Kuzborska, 2011; Niu & Andrews, 2012; Li, 2013). Most of these researchers have focused on the relationship between teacher beliefs and their instructional practices. Some of the findings of studies related to the aforesaid relationship are useful for the current research. They are:

- Teacher knowledge and beliefs are intricately woven and related.
- Beliefs are like a watch dog that monitors and guides acceptance, rejection, processing and use of new information.
- Teachers' beliefs are both personal and social in nature and affect their behaviour.
- Teacher beliefs can be held both consciously and unconsciously and thus be captured from observation of external behaviour, statements of teachers and intentions of teachers.
- Classroom practices also shape teacher beliefs.
- It takes a lot of time to change teacher beliefs.
- There is very little known about how to measure and judge beliefs.

3.5.1 Beliefs of In-service Teachers

There is plenty of research on beliefs and attitudes in the field of second/foreign language teacher education (Johnson, 1992; Scott and Rodgers, 1995; Lamie, 2004; Liu & Fisher, 2006; Busch, 2010; Borg, 2011). But studies concentrating on in-service teachers' beliefs are much less in number when compared to those on pre-service teachers' (Borg, 2011). Moreover, there is a great deal of contradiction in the findings of the studies related to the impact of in-service teacher beliefs on their practices. Some of the studies are presented in the following table:

Table 3.1: Research on the relationship between teachers' beliefs and teaching practices

Researcher(s) and year of publication	Aim of the study	Findings
Johnson (1992)	To determine the extent to which ESL teachers' instructional practice is consistent with their theoretical beliefs	Teachers' instructional practices are to a great extent (60%) consistent with their beliefs.
Farrell & Particia (2005)	To compare the beliefs and practices of ESL teachers	Teachers' practices are not consistent with their beliefs.
Lorduy, Lambraño, Garcés & Bejarano (2009)	To examine to what extent English (EFL) teachers' teaching practice is related to their beliefs	Teachers' teaching practices are not consistent with their beliefs.
Khonamri & Salimi (2010)	To examine to what extent English (EFL) teachers' teaching practices are consistent with their beliefs	There is inconsistency between teachers' beliefs and their teaching practices.

Researchers have tried to explain the discrepancies in the findings by pointing out that, sociocultural, personal and psychological factors often affect teachers' beliefs about teaching and learning. Similar kind of divergences can also be found in studies that have tried to measure the impact of in-service teacher education/development programmes on teacher beliefs. A summary of some of the important studies is presented below.

Table 3.2 Research on the impact of in-service programmes on teachers' beliefs

Researcher(s) and year of publication	Aim of the study	Findings
Scott & Rodgers (1995)	To measure the impact of in-service TD programme on their attitude	To a great extent, the programme had an impact on teachers' attitudes.
Pacek (1996)	To measure the impact of an in-service programme on EFL teachers' beliefs	Many participants found it difficult to change their beliefs and attitudes at the end.
Lamie (2001)	To measure the impact of in-service on EFL teacher beliefs	Participants are divided in terms of their beliefs about the programme inputs.
Kurihara & Samimy (2007)	To capture the impact of in-service training on EFL teacher beliefs	Participants showed positive attitude towards training but found it difficult to apply the new knowledge.
Phipps (2010)	To measure the impact of in-service training on an English language teacher's beliefs	The teacher showed change in belief and beliefs got strengthened.
Borg (2011)	To assess the impact of an in-service programme on English language teachers' beliefs	Some change in beliefs is observed. More changes could have been brought about.

The above-mentioned reviews very clearly indicate the lack of consistency in the impact of INSET programmes on teachers' beliefs. According to Hall (2005), teacher beliefs are difficult to change, but she also agrees that they need to be challenged creatively for triggering such changes. The review also indicates that there is a need to look into the process of change in beliefs as very few studies have focused on the actual process of change.

Since any effective teacher education programme is based on a set of guiding principles, the existing research literature in this area is reviewed in the next section to arrive at a set of principles for the current study.

3.6 Guiding Principles of an Effective TD Programme for In-service Teachers

Wallace (1991) proposes three models of teacher professional development: the craft model, the applied science model and the reflective model. According to the first one, as the word 'craft' suggests, the teacher has to learn the craft of teaching from the teacher educator, the master craftsman, through careful observation and imitative practice. In case of the second model, the teacher is made aware of the already-proved effective teaching practices and is expected to translate the research-based knowledge to his/her own practices. In contrast, the last model emphasizes experiential learning in which the teacher reflects on his/her experience of learning before applying the new learning in their own context. Though researchers often emphasize the importance of reflection in TD, the first two models can also be wisely used for developing effective TD programmes for in-service language teachers. To avoid the possible dangers of an overstretched reflective model, Ur (1999) suggests 'enriched reflection'— a flexible reflective model that accommodates a balanced amount of external input.

Creemers, Kyriakides and Antoniou (2013) analyse research related to the competency-based approach (CBA) to TD, which was popular in 1970s in the US and 1980s in Europe, and that based on reflective practice, which is the most dominant approach at present, and list the strengths and weaknesses of the two approaches. The following table contains the details:

CBA	Reflective Approach
Strengths: • has well-defined and specific objectives • can lead to change • provides more clarity about target skills • is good for small number of goals • focuses on specific competencies required by teachers • provides hands-on experience	Strengths: • is helpful in enabling teachers to analyse and improve their practice • brings flexibility to beliefs about teaching • promotes autonomy and critical thinking • integrates teachers' existing knowledge and experience • provides opportunities for theory-building
Weaknesses: • lacks focus on overall competencies • adopts a discrete point approach to learning individual skills • is mechanical in dividing the profession into skills • puts very little emphasis on critical and creative thinking • pays little attention to theoretical knowledge and principles of teaching • does not support teacher independence • ignores personal knowledge	Weaknesses: • ignores research-based findings about the necessity of specific skills • is not often systematic in making teachers reflect • is vague in terms of content • does not use foolproof tools for reflection • neglects practice • does not pay much attention to change and impact

(pp. 23-28, 37-42)

It is evident from the review of research literature presented in the previous sections that for a TD programme to be effective, the strengths of different models should be brought together and utilised as the guiding principles for conducting the programme. So, an effective TD course should:

• have a clearly-defined and specific focus

• address participants' professional needs

• take into account participants' existing knowledge, belief, experience and institutional policies

• be experiential in nature

- have a balanced mixture of external input and individual reflection
- enable participants to become independent practitioners

3.7 Short TD Courses for In-service Teachers

Short TD courses are quite common and teachers attend more short courses than long programmes. These courses often have a narrow focus in terms of skills and knowledge. Considering that this study included a short TD course, it is pertinent to review some of the research findings related to the drawbacks and benefits of such courses.

3.7.1 Drawbacks

After systematically analysing short in-service TD programmes, Tomlinson (1988) claims that "participants will lose more than they will gain" (p. 19) after participating in such programmes. He cites several reasons which include teacher educators' lack of knowledge about teachers' teaching context, their sense of superiority and disregard for teachers' existing knowledge, highly theoretical nature of programmes, little opportunity for teachers to practise, rigid nature of course components, overloading of information, impractical objectives, no support for teachers after the programme, etc. Lamb (1995) discusses a similar experience and points out that the short courses created confusion among teachers and they had little impact on teachers' beliefs. He believes that teachers take time to internalize and assimilate new ideas, and the short course did not give them that opportunity. Often, short courses are not 'adequately spaced' (Guskey, 2000, p. 23). As a result, teachers find little time to internalize the new learning and utilize it in their own context (Waters & Vilches, 2000). Rodrigues (2005) calls such courses "bite size pieces of teacher professional development" (p. 6) and feels the need for longer and more continual effort at helping teachers professionally. But she adds that such programmes are too high in number and quite wide-spread. Sim (2011) feels that one of the strong reasons behind the failure of many short courses is the participants' lack of interest in attending them.

3.7.2 Benefits

Miles (1964) and Ruddock (1981) discuss the advantages of short courses. Rudduck's views are very relevant. She believes that teachers can get plenty of new ideas, share their professional experience, avoid feeling professionally segregated and choose what interests them. Wolter (2000) asserts that a participant-centred approach to course design

can help construct a good short in-service TD programme. A thorough needs analysis should be carried out to obtain adequate information about the participants' needs.

Waters and Vilches (2000) analyse a short TD programme for in-service teachers in Philippines and report that such a course should be "less of an end in itself, and much more a vehicle for fostering learning by teachers within their normal cultural milieu" (p. 133) to be effective. Thus, while designing short courses, the course designers and teacher educators should make the participants aware of the possibilities of change and improvement in their practice and let them choose and use what they can in their respective school situations.

In addition to the above mentioned benefits, short courses are easy to construct and economical in nature, can address the professional needs of specific groups of participating teachers and can be evaluated and improved with less effort. Sim (2011) points out that the critics of short courses often expect such courses to demonstrate as much impact as longer programmes and "oversimplify the complexity of teacher practices within an institution" (p. 46).

Tomlinson (1988) sets a few ground rules for the creation of effective short courses. According to him, a short course should have a very limited and specific focus; the teacher educator should be familiar with the participants' teaching context; it should utilize the participants' existing knowledge; and the course should provide hands-on experience to the participants.

3.8 Teacher Change

Most discussions on teachers' knowledge, belief, attitude and practices in language education are related to teacher education in general and refining and improving teachers' instructional practice in particular. Such discussions ultimately help in developing and fine-tuning different kinds of teacher education and development programmes. Though these programmes may vary from one educational context to another and one group of teachers to another in terms of content and the way they are imparted, yet most of them, as observed by Guskey (2002), "are systematic efforts to bring about change in the classroom practices of teachers, in their attitudes and beliefs, and in the learning outcomes of students" (p. 381). According to Richards, Gallo and Renandya (2001), "change can refer to many things including knowledge, beliefs, attitudes, understanding,

self-awareness, and teaching practices" (p. 1). So TD programmes, in the best possible scenario, should be able to trigger positive change in teachers.

Smith, Hofer, Gillespie, Solomon and Rowe (2003) in a historical review of research on change as envisaged in professional development programmes point out that in the 1960s and 1970s, changing teacher behaviour; in the 1980s, supporting school reform; in the early 1990s, improving student learning; and in the late 1990s, enhancing student quality were the goals (p. 6). The review does not cover research on teacher as an individual with a set of beliefs, attitudes and perceptions of his/her own. In a more recent review, Kubanyiova (2012) claims that there have been two approaches to studying teacher change: the first one "has examined teacher change within the broader social, cultural and political contexts of school organization" (p. 6) and the second approach "has focused on individual or small group change and has investigated cognitive, affective and behavioural change processes in teachers" (p. 6). The second perspective has become more popular in recent times considering the fact that there is no conclusive research about how changes get reflected in teaching practices.

3.8.1 Perspectives on Teacher Change

Several models have been proposed for teacher change. These models can be viewed from two broad perspectives as suggested by Richardson and Placier (2001):

- empirical-rational
- normative-re-educative

From the first perspective, teachers are expected to follow the ways or approaches suggested to them in a training programme. This top-down approach to promoting teacher change is 'mandated' (Stivers & Cramer, 2009). As the change is primarily initiated from outside, there can be resistance from teachers' side and it can put teachers in stressful situations. But Sim (2011) looks at resistance as a natural part of change and asserts that innovation is not generally accepted by teachers easily. Lamie (2002) and Waters (2006) have also discussed how innovation and change are intricately connected and how these factors need attention in INSET programmes for language teachers. Fullan and Hargreaves (1992) too talk about resistance in detail.

As opposed to the empirical-rational perspective, the normative-re-educative perspective looks at teacher change as something that is 'voluntary' or 'naturalistic' (Kubanyiova,

2012, p. 7) emerging out of teachers' professional practice, participation in TD activities and conversation and collaboration with colleagues. This perspective marks a departure from the behaviouristic approach to studying teacher change to a cognitive one. This is experiential in nature and can form a better foundation to INSET or TD programmes (Kolb, 1993; Crandall, 1994; Waters, 2006). However, it is important that INSET programmes are linked to teachers' actual practice for a longer period of time (Guskey, 2002).

Stivers and Cramer (2009) discuss two other types of teacher change: incremental and fundamental (pp. 31- 33). The first of these refers to changes that teachers bring in to improve existing practices without necessarily changing anything substantially. On the other hand, fundamental changes (Cuban, 1992) come to play when teachers strongly feel that there is something really wrong with the existing practice and thus, a complete overhaul is required. Fundamental changes happen very slowly and are not often observed immediately after INSET programmes.

As mentioned earlier in this section, research on teacher change is now closer to the area of teacher cognition than ever before. The effect of INSET on teachers and their course of change have also been discussed in teacher cognition research (Borg, 2006) and from sociocultural perspectives (Johnson & Golombek, 2011). Though these are formally accepted as two domains, they are closely interrelated. With sociocultural learning theories dominating language education research, teachers' cognition, i. e., "belief, knowledge and thoughts" (Borg, 2009, p. 166), is being studied in relation to the educational and sociocultural context of teachers' professional practice. Thus, teacher change from this perspective may mean change in teachers' knowledge, beliefs and attitude. It implies that teacher change must consider the interplay of factors like "language teacher education", "teacher learning", "teaching practice" and "students' learning" (Kubanyiova, 2012, p. 25). This interplay or "relationship of influence" (Freeman & Johnson, 2005, pp. 76–80) needs more exploration for any reliable conclusion to be reached.

3.8.2 Strategies to Activate Change

There are three types of strategies suggested by Chin and Benne (1969) for implementing change:

- Power-coercive strategy

- Rational-empirical strategy

- Normative-re-educative strategy

(as cited in Lamie, 2005, pp. 17-22)

As the names suggest, in case of the first one, political, economic and moral powers are applied to enforce change; the next one depends on using information to infuse change; and the last one focuses on changing the culture of schools to involve the teacher in the process of change. The first strategy has been considered as a failure since it does not entertain the teacher's individual way of thinking (Lamie, 2001). The second strategy adopts a rational approach which is based on the premise that teachers will accept the change if its benefits are explained to them logically. But this kind of argument takes it for granted that everyone thinks rationally, which is problematic (Fullan, 1991). The last one is better than the first two in that, here the course of change is determined collaboratively by all the stakeholders. But it is time-consuming and difficult to implement in a large-scale set-up.

3.8.3 Classification of Teacher Change

Before getting into classification of teacher change, some clarification regarding the scope of classification is necessary. Here, the classification has nothing to with the areas like 'knowledge', 'belief', 'concept' and 'attitude' directly. The concern is limited to the kinds of changes observed after an effort is made to keep track of them. Smith et al. (2003) have proposed a classification of teacher change into what they call "three-category spectrum of change" (p. 13). The three categories are:

- no change
- non-integrated change (thinking or acting changes)
- integrated change   (p. 13)

The category 'no change' refers to very little or no change, 'non-integrated' has to do with teachers' intention to accept changes and the third category is all about translating proposed changes into practice. Although, these categories are not absolute in any sense, they nonetheless provide some kind of clarity to understanding teacher change effected by any TD programme.

3.8.4 Gaps in Research on Teacher Change

There has been a lot of research on the impact of in-service training programmes on teacher performance. But there is virtually no study about the impact of training in assessment on teacher performance and teacher change. Even fewer attempts have been made to look into how contextual factors have impact on the change and development of individual teachers (Zeichner & Gore, 1990). Another area that needs research attention is the impact of professional development programmes on teacher belief systems and how the change in beliefs gets reflected in practice (Tatto, 1998). Also, the lack of methodological innovations in studies on teacher change needs to be addressed (Kubanyiova, 2012). More attempts should be made to focus on individual teacher change through multiple case studies (Grossman, 2005). Kubanyiova (2012) suggests that mixed methods approaches can be productive in capturing teacher change.

3.8.5 An Eclectic Framework for Understanding Teacher Change

Kubanyiova (2012) proposes an integrated model of Language Teacher Conceptual Change after defining the impact of teacher education, cognitive engagement of teachers at a deeper level, effect of affective and motivational factors and teachers' possible selves. This is indeed a useful model as it takes into account teachers' existing knowledge, beliefs, attitudes towards aspects of language education, sociocultural realities in which the teacher's instructional practices are situated, the curricular demands on the teacher, etc. But the irrefutably dynamic nature of teacher change needs a little more attention.

According to Kumaravadivelu (2001), any transmission model of teacher education is "hopelessly inadequate" (p. 552) in helping teachers become "self-directing and self-determining" (p. 552). Knowledge-construction as suggested by Johnson (2009) has thus gained in popularity as it emphasizes co-construction of teachers' knowledge-base with the help of available input. Furthermore, if the dynamicity of curricular requirements are believed and considered, then teacher change must be accepted as a dynamic process. The following spirally-progressive framework is quite inclusive and flexible and offers a wider scope to understand teacher change:
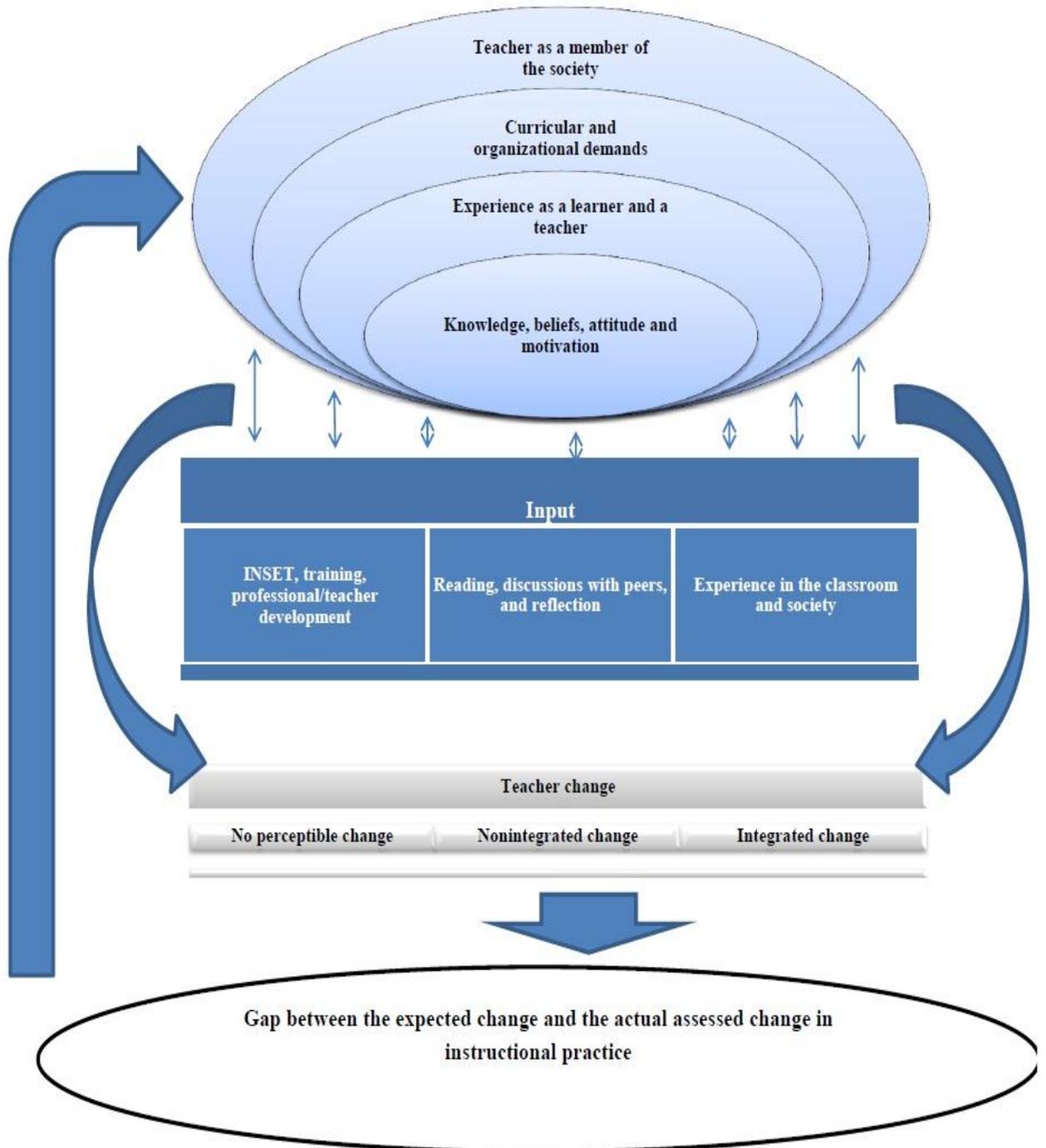
Figure 3.1: Working framework to understand teacher change

As shown in the above framework, it is believed that teachers' existing experience, knowledge, beliefs, etc. interact with the input provided in the form of INSET programmes, reading, discussions with peers, etc. During this interaction, the teacher, a dynamic socio-cognitive entity, chooses to do what they think personally is the most appropriate thing to do in the given circumstances. The decisions regarding this include,

but are certainly not confined to, what to accept, what to reject, how much of the input is to be accepted, how much to be rejected, how much to be included in the instructional practice and a possible time-frame for putting the accepted components into practice. After the change is assessed and the gap between the expected change and the actual measured change in instructional practice is discovered, the process of change continues and follows a path similar to the one shown in the framework. However, the position of each stage also gets changed with time and hence, a set of new definitions gets operationalized in the process.

The forthcoming sections concentrate on research in the field of assessment education in teacher education programmes, which is also the area under scrutiny in the current study.

3.9 Assessment in Teacher Education Programmes

Efforts to train teachers in assessment started gaining attention almost fifty years ago. Conant (1963) in his book *The Education of American Teachers* and a few others like Mayo (1964), Goslin (1967), etc. drew attention to the need of training teachers in different aspects of educational measurement. Gullickson (1993) makes a thorough review of these early efforts and suggests amendments in the course content of teacher education programmes. Fifty years after Conant and twenty years after Gullickson, we are still grappling with the same problem. Stiggins (1993, 2007) observes that very few teacher-preparation programmes across the world focus on developing teacher knowledge about assessment, and even in in-service TD programmes, assessment has traditionally been one of the neglected areas. He goes on to point out some of the possible reasons of neglect. They include the following:

- Our system of educational management is more process-based than outcome-based.
- Assessment courses are academically more challenging and demanding than other teacher education courses.
- Assessment training may pave the way for accurate assessment of student achievement which may risk the reputation of many educational institutes.
- There is a wide-spread belief that externally-designed assessments are better than teacher-made assessments.

(Stiggins, 1993, pp. 29-30)

The above reasons are quite convincing. In the Indian context, however, there could be more to the neglect of assessment training in teacher preparation programmes than the aforesaid reasons. The scarcity of experts in assessment and lack of assessment literacy among other teacher trainers could be some of the possible reasons. The fact is that the lack or sometimes absence of training in assessment for teachers is hurting the education systems across the world, and teacher education programmes must make efforts to equip teachers with adequate knowledge and skills so that teachers can assess their students in their respective classrooms. In the next two sections, assessment training in both pre- and in-service teacher education programmes with special reference to ELE is reviewed.

3.9.1 Assessment Training in Teacher Preparation Programmes across the World

Researchers working in the area of teacher training/development in assessment (Stiggins, 1993; Malone, 2007; Davies, 2008; Fulcher, 2012; Stoynoff & Coombe, 2012; and Leung, 2013) have found inadequacies with assessment training in teacher preparation programmes. All the above-mentioned researchers and many others have discussed in detail how teachers' participation in in-service TD programmes that exclusively address teachers' assessment needs help them with their assessment practices. Before examining the in-service programmes, it may be apt to review some of the reports on pre-service programmes across the world.

In its report released in 2012, the National Centre on Teacher Quality (NCTQ) declared that 21% of the sampled teacher education programmes in the US train teachers in the basics of assessment, only 1% of the programmes teach teachers how to analyse assessments, and less than 2% prepares teachers for using assessment results for making instructional decisions. Overall, only 3% of the teacher education programmes pay satisfactory attention to assessment. Earlier in 1999, Stiggins conducted a similar survey. The survey revealed that "only 25 of the 50 states require that teachers either meet specific competence standards or at least complete assessment coursework during their preparation" (p. 23).

Bachman (2000) claimed that only half of the TESOL members (during 1990s) had a course in language testing. This provides a broader picture of the scenario under examination because TESOL members include teachers, teacher trainers, university teachers, etc. and most of these people are expected to have access to courses in language testing. Seven years later, Stoynoff (2007) came up with very similar findings. He found

that only 50% of the TESOL teacher education programmes offered courses in language assessment.

In Europe, the situation is not very different from what is described above. In a relatively recent large-scale survey of language teachers (mostly English) practicing across Europe, Vogt et al. (2008) discovered that 28% of teachers had no training in assessment, 35% got only a little training in preparing classroom tests, 40% of the teachers had no idea about using ready-made tests, 31% reported having no training and 37%, some training in providing feedback to students. Vogt et al. also reported that teachers have very little training in using portfolio, self- and peer-assessment, and informal and formative assessments.

The inattention to language assessment courses in teacher education programmes is not limited to the US or Europe. Brindley (2001) and O'Loughlin (2006) have discussed dearth of training for teachers in Australia, Falvey and Cheng (1995) and Qian (2008) observe a similar plight of teachers in Hong Kong, Shohamy (1998) mentions that teachers in Israel have very limited knowledge about carrying out classroom assessments efficiently, and Troudi, Coombe and Al-Hamley (2009) discover that teachers in Saudi Arabia and Kuwait are aware of their lack of knowledge about assessment.

The brief overview in this section presents a rather dismal picture. The belief that "traditionally teacher education has generally not given assessment literacy – that is, the professional knowledge and repertoire regarding assessment – a great deal of curriculum prominence" (Leung, 2013, p. 1510) has been confirmed yet again. On the other hand, many researchers have suggested bridging the gap between assessment demands on teachers and the shortage in supply of professional support or necessary training in assessment for them through different TD programmes. This aspect is discussed below.

3.9.2 TD in Language Assessment

What is meant here by "TD" is very close to what Mann (2005) meant by the same term in his state-of-the-art article on language teachers' development. It is important to define it because the teacher as well as teaching, both are socioculturally placed and any kind of development cannot be complete without the teacher willing to play "an active role in their own development processes" (Mann, 2005, p. 104). In the light of this definition, some of the suggestions for developing in-service teachers' language assessment literacy are considered below.

It is true that a teacher cannot function effectively in the classroom without the knowledge of assessment, but, as Brindley (1997) suggests, it may not be realistic to expect teachers to design good assessment tasks for their students in the absence of effective training or TD programmes in language assessment. Developing language teachers' professional knowledge about assessment, popularly known as 'assessment literacy' in language assessment literature, has generated some research interest in the last two decades (Stiggins, 1991; Brindley, 1997; Inbar-Lourie, 2008; Taylor, 2009; Fulcher, 2012; etc.). Though there is agreement among experts about the need of language assessment literacy for English teachers, there have been debates on what it should comprise. Again, the impact of assessment literacy on the assessment skills and knowledge is yet to be empirically supported in ELE research.

The most recent review that analyses and systematically presents resources of TD is by Stoynoff and Coombe (2012). This work sounds more convincing and comprehensive than two earlier attempts by Richards and Farrell (2005) and Malone (2008). Richards and Farrell (2005) looked at such resources from 'individual' and 'institutional' perspectives, whereas Malone (2008) divided those into 'text-based' and 'technology-mediated' materials. Stoynoff and Coombe (2012) take into account the above-mentioned views and classify the resources under "medium (print and Web-based), sponsors (professional associations, nonprofit organizations, and government-supported entities), and teacher-based activities" (p. 124) while concentrating on TD in language assessment. The following division of resources is broadly based on the one suggested by Stoynoff and Coombe.

Print and Web-Based Resources

Davies (2008) examines language assessment textbooks used by language educators since Lado's times. But he does not discuss the suitability of the popular textbooks for school teachers in ESL and TESOL contexts. The applicability of language assessment textbooks in these contexts needs to be empirically examined before accepting them as realistic sources of TD. However, some articles and research papers focusing on assessment of particular language skills or components, use of authentic assessment methods, construction of classroom assessments, etc. can be used by teachers. Again, the difficulty level of the materials and the help teachers need to understand and make use of these materials should be considered realistically.

Web-based materials, in contrast to the above, are much more flexible and user-friendly. Teachers have an array of options to choose from. They can easily find what they want and in the process, learn to tailor and adapt the available materials for their own students. The following are some of the popular websites that provide information about language assessment:

- http://languagetesting.info/
- http://www.cal.org/calwebdb/flad/
- http://www.nclrc.org/essentials/index.htm

There are also websites exclusively dedicated to assessment of individual language skills and components. And apart from these, webinars, virtual workshops, online TD courses, etc. can also be used by teachers for professional development. But development of assessment skills through these resources is dependent on factors like access to internet, basic idea about available resources and above all, motivation to grow as a professional.

Resources Accessed through Organizational Support

Local, national and international teachers' associations often have special interest groups, conferences, seminars, etc. focusing on language assessment. Sometimes the school/institution where the teacher works and at other times, the government take the initiative to sponsor programmes for developing assessment skills of teachers. Even peer-based discussions can be of great help in this regard. However, any attempt to 'train' teachers and thereby impose a fixed set of ways on their functioning may not yield the desired result (Mann, 2005). Bartels (2005) shows that teachers fail to apply knowledge that is presented to them in a top-down manner. On the contrary, a top-down approach grounded in the sociocultural realities in which the teacher practices (Richards & Farrell, 2005; Richards, 2008; Brown, 2008) can be more effective when it comes to learning to assess effectively.

Teacher-Based Activities

Stoynoff and Coombe (2012) do not mention clearly what they mean by 'teacher-based activities'. But it may possibly mean what a teacher can do at individual level. Reflection on the process of assessment, students' performance in the classroom, their individual development of language ability, items used in the assessment, the feedback provided to students, the impact of assessment on students and the assessment policies, and even

other activities like keeping journals, maintaining a diary, etc. that are initiated by the teacher him/herself can be included under this category. Richard and Farrell (2005) believe that teacher reflection can be supported and guided by the institution the teacher works in.

The three types of resources suggested by Stoynoff and Coombe overlap with each other in their scope and functioning. Though this kind of division provides certain amount of clarity, there is a fear that this might narrow the strategic platform and the research base on which in-service development of teachers is planned and executed. The matter becomes further complex when the focus is on language assessment. The dearth or absence of pre-service training in assessment may require a very different approach to in-service development of assessment ability of teachers. Their beliefs, the assessment policy, the general pedagogic context, the place of the English language in the context, the teacher education policy, etc. need to be analysed and considered before finalising the components of assessment and the resources through which teachers can learn to use these components.

The above discussion calls for serious thinking about providing proper training to teachers before asking them to assess their students on their own. In India, the recent Examination Reforms (2006) has made some "structural and procedural change" to the manner in which assessment was being practiced in schools. These changes have initiated de-centralization of assessment in the country by entrusting the responsibility of assessing the language skills of students to their respective English teachers. However, very little professional support has been provided to develop teachers' professional knowledge in assessment. The current research project tries to look into this issue and made a small plea for providing assessment training to teachers.

As the current study concerns the impact of in-service development of assessment literacy of teachers, it is necessary to have a brief review of research on in-service education of language teachers and its impact on them, and teacher knowledge, beliefs and change with special reference to language assessment.

3.9.3 Language Teachers' Beliefs about Assessment

There is a complex relationship among what teachers say, what they do and what they believe they should do. Research on teacher beliefs has not been able to arrive at any pattern in teachers' beliefs. This lack of pattern is evident in teachers' beliefs regarding

different aspects of language education. In the following section, a fresh examination of research literature on language teachers' beliefs about assessment is presented.

According to Yang (2007), there are two main elements of teachers' beliefs about classroom assessment:

- Teacher beliefs about the pedagogical benefits of classroom assessment
- Teacher beliefs about the difficulty of implementing classroom assessment

Some of the important studies on second/foreign language teacher beliefs about assessment practices are laid out briefly in the following table:

Table 3.3 Research on the relationship between teachers' beliefs and assessment practices

Name(s) of the author(s), year of publication and research method(s)	Aim of the study	Findings
Cox (1994), survey	To study teachers' classroom assessment practices	There is no consistent relationship between teachers' beliefs and their assessment practices.
Bliem & Davinroy (1997), oral reports & standardized interview	To find out if teachers' assessment practices are influenced by their existing beliefs about assessment	Teachers view and interpret new assessment practices through the lens of their beliefs.
McMillan (2003), qualitative survey & interview	To explore the relationship between teachers' beliefs and their assessment practices	The beliefs do not always indicate teachers' actual assessment practices.
Yang (2007), quantitative survey	To explore the extent to which teacher beliefs, training in assessment and teachers' reports about their competency in classroom assessment are related to classroom assessment practices	There exists a strong positive relationship between what teachers believe about alternative assessment and what they practice.
Brown & Harris (2010), non-experimental survey and interview	To study the relationship between teacher conceptions of assessment and the practices of classroom assessment	There is no significant relationship between the two factors.
Bullock (2010), attitude questionnaire & standardized open-ended interview	To find out about the relationship between teachers' attitudes, beliefs, and practice with regard to self-assessment	Teachers' behavioural beliefs have positive impact on their intentions.
Muñoz, Palacio & Escobar (2012), running records & standardized interviews	To study how teachers' beliefs about assessment get reflected in how they assess	There are contradictions between what teachers say they do and what they believe.

It is very obvious from the research findings presented in the above table that teacher beliefs do not necessarily determine their assessment practices. Only Bullock's (2010) research suggests that behavioural beliefs are better indicators of teachers' assessment

practices. However, most of the researchers who have worked on language teachers' beliefs about assessment agree that teachers' beliefs and assessment practices do not have a straightforward relationship because contextual factors play a major role. Davison (2004) mentions that "less attention has been paid to the way in which different educational and cultural contexts, and teacher assumptions about those contexts, shape teachers' assessment beliefs, attitudes and practices" (p. 306). Thus, it is important that teacher education programmes do not overemphasize the change in teachers' beliefs while assessing the impact of the programme on teachers' assessment practices. Whether research on teacher change in language education has led to similar kind of view regarding teacher belief is something that forms a part of the discussion in the next section.

3.10 Changing Assessment Practice through TD Programmes

If the above framework is applied to developing a TD programme in assessment with a view to effecting some positive changes in their practice, then certain relevant things need to be made clear in the light of findings of two major studies: one by Chinda (2009) and the other by Jeong (2011).

In his study, Chinda (2009) found that any professional development (PD) programme aiming to improve teachers' assessment practice must "match what teachers do and already know in assessment" (p. 254). He also proposes "a rigorous background study of teachers' needs in that particular context" (p. 254) so that the programme can have positive impact on teachers. Chinda also suggests that the PD needs to provide hands-on experience to teachers to maximize the impact. The variety in teachers' experience combined with an informal approach to the implementation of the programme can yield rich outcomes.

Jeong's findings are very similar to those of Chinda. Jeong points out that the programme instructor should be aware of the needs of teacher-participants, value their existing experience as teachers, and build these elements into the course. Like Chinda, she also emphasizes providing practical and relevant experience in assessment to teachers during the programme.

The findings of studies by Chinda (2009) and Jeong (2011) can be integrated into a single comprehensive framework for understanding teacher change. This resultant framework

can be used for planning, designing and carrying out TD programmes for in-service teachers.

As one of the main concerns of the current study is improving assessment practices of in-service teachers by developing their classroom language assessment literacy through a TD programme, the proposed comprehensive framework presented below tries to meet that end.

Stage 1	• *Collection of information about teachers' previous experience as a teacher and training, current assessment practices and personal belief and motivation in relation to it, assessment policy, organizational demands and level of CLAL*
Stage 2	• *Analysis of the above data and teachers' classroom language assessment needs*
Stage 3	• *Development of a teacher development programme in classroom assessment in consultation with the participating-teachers, a few local teacher trainers and relevant research findings*
Stage 4	• *Transaction of the programme in an informal manner and with the help of practical tasks that offer ample scope to teachers to work on those and improve their level of CLAL*
Stage 5	• *Follow up by collecting information about how teachers are planning to implement or are implementing new learning and helping them with it*
Stage 6	• *Collection of feedback from teachers about the programme, evaluation of the programme and change in teachers' classroom assessment practices prompted by the programme and incorporation of changes, if necessary, to the programme*

Figure 3.2: Framework for TD in language assessment

The above framework is cyclical and flexible in nature. So Stage 6 can be followed by Stage 1 and a few more components can be added to each stage of the programme if required. In fact, the proposed stages are not watertight compartments; they have been created to add precision and clarity to the planning, construction and transaction of such programmes. Each stage should be accepted as a construct and can therefore be further

defined, elaborated and then operationalized as per the requirement of individual educational contexts.

3.11 Conclusion

This chapter attempts to review and present research on TD and change with special reference to language assessment. In the process, in-service language teacher education, teacher knowledge and teacher beliefs have been discussed as part of the larger areas of focus. It must be made clear that the research literature examined and used for development of the theoretical framework for understanding change is limited in its scope and only relevant details pertaining to the problem under scrutiny have been incorporated in the review. In addition, no claim has been made about the nature of TD programmes that can lead to the desired change in teachers' assessment practices. The suggested theoretical framework offers only a flexible research-driven platform to study the impact of assessment literacy on teachers' assessment practices and the relevance of TD programmes in this connection.

CHAPTER 4

RESEARCH METHODOLOGY AND PROCEDURE OF DATA COLLECTION

4.1 Introduction

This chapter discusses in detail the research design and methodology and the process of data collection followed in the study. In the first part of the chapter, the purpose of the study, the hypotheses on which this study was based and the research questions that guided the study are highlighted. Then there is a discussion on the Mixed Methods Approach, which includes its definition and the rationale for its use in the current study. This is followed by individual sections focusing on the sampling design and the rationale for its selection, description of cases, methods of data collection and rationale for their employment and the procedure of data collection.

4.2 Purpose of the Study

The study tried to examine the relationship between teachers' CLAL and classroom assessment practices. It also made an attempt to trace the impact of a need-based TD programme in classroom assessment of language ability on the participant-teachers' CLAL and subsequently, their assessment practices which include ability to design assessment tasks and criteria, offer feedback and appropriate beliefs about assessment. The study was based on an informed assumption that the practice of examination-centred teaching can be changed into teaching-integrated assessment if teachers' CLAL is developed.

4.3 Hypotheses

The study was based on the following hypotheses which were specific to the context, i. e., Andhra Pradesh, in which the study was carried out:

- There is a relationship between teachers' level of assessment literacy and language assessment practices.
- A need-based TD programme in CA of language ability will lead to change in their assessment practices.

4.4 Research Questions

The current study tried to find answers mainly to the following questions:

- What is the average CLAL level of secondary school English teachers in the state?
- How is teachers' CLAL related to their assessment practices?
- What impact does a short TD programme in CA of language ability have on the teachers' ability to design classroom assessment tasks and assessment criteria, provide feedback, and on the teachers' beliefs about assessment?
- How do teachers respond to the programme?

4.5 Mixed Methods Approach

Mixed methods research has been defined in *The SAGE Dictionary of Social Research Methods* as the "combined use of both quantitative and qualitative methodologies within the same study in order to address a single research question" (Jupp, 2006, p. 179). The mixing of qualitative and quantitative methods happens in almost all stages of a study, i.e., from construction and planning of the study to data collection, analysis and interpretation. Philosophically grounded in pragmatism, mixed methods approach includes features of the qualitative tradition which is based on constructivism and the quantitative tradition which is positivist in its orientation. The argument is that it is wise and logical to find the best possible way that can guide in obtaining answers to research questions rather than rigidly adhering to the normative boundaries of any one particular approach. In fact, the 'hybrid vigour' (Miles & Huberman, 1994, p. 310) offered by this eclectic approach comes along with a lot of flexibility which helps the researcher in addressing the research questions more effectively. In other words, there is always room for varying proportions of quantitative and qualitative ways depending on the research problem.

4.5.1 Convergent Parallel Design

Creswell and Plano-Clark (2010) call Convergent Parallel Design as the "most well-known approach to mixing methods" (p. 77). In this kind of design, qualitative and quantitative data are collected and analysed separately though concurrently. Both forms of data collection get equal importance. Then findings obtained through both the strands are compared and corroborated to arrive at meta-inferences.

Convergent parallel design is an effective way to strengthen the study by combining the advantages of qualitative and quantitative strands and overcome the inherent weaknesses of both. As pointed out by Dornyei (2007), qualitative research is often targeted for being "too context-specific and employing unrepresentative samples" (p. 45) and quantitative research is often viewed "as overly simplistic, decontextualized and reductionist in terms of its generalizations" (p. 45). The complementary nature of the obtained data may further lead to improved validity.

For the current study, however, an adapted version of the convergent parallel design was adopted. This kind of adaptation is discussed by Creswell and Plano-Clark (2010, p. 55). Here, the convergent parallel design started like a sequential exploratory design in which qualitative procedures of data collection followed a quantitative one. This was done to overcome the problems associated with parallel collection of data. In addition, this variation also helped in choosing an appropriate sampling strategy for the qualitative stage.

The following diagram shows how it was used for this study:

```
                  ┌─────────────────────────────────────────┐
                  │  Quantitative data collection and analysis │
                  └─────────────────────────────────────────┘
                                      │
                                      ▼
┌──────────────┐        ┌──────────────────┐        ┌──────────────┐
│  Qualitative │ ◄───── │ Data collection  │ ─────► │ Quantitative │
└──────────────┘        │   and analysis   │        └──────────────┘
                        └──────────────────┘
                                  │
                                  ▼
                          ┌──────────────┐
                          │ Intervention │
                          └──────────────┘
                                  │
                                  ▼
┌──────────────┐        ┌──────────────────┐        ┌──────────────┐
│  Qualitative │ ◄───── │ Data collection  │ ─────► │ Quantitative │
└──────────────┘        │   and analysis   │        └──────────────┘
                        └──────────────────┘
                                  │
                                  ▼
               ┌─────────────────────────────────────┐
               │  Meta-inference and interpretation  │
               └─────────────────────────────────────┘
```
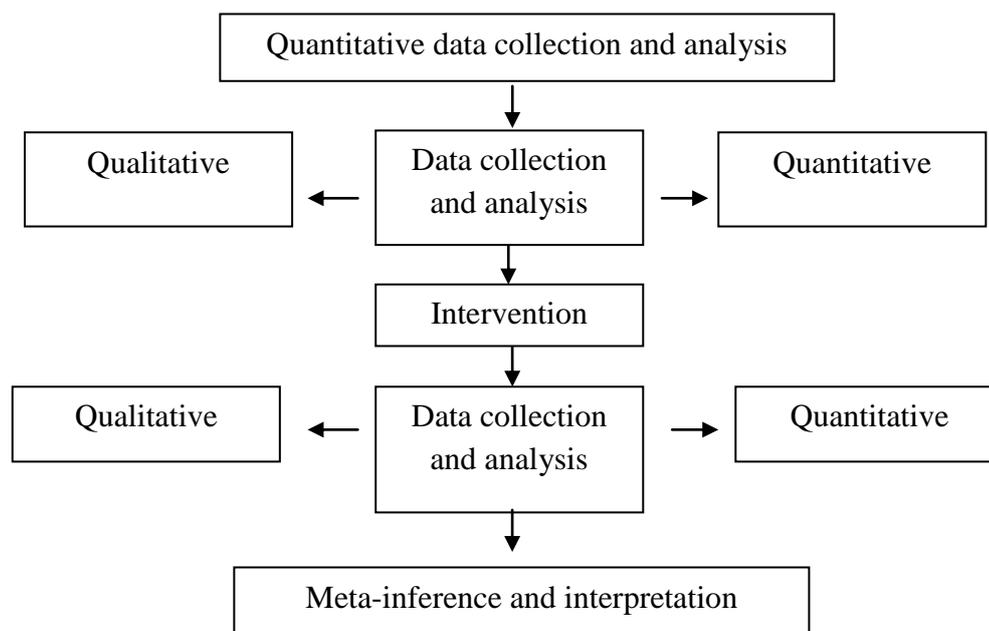
Figure 4.1: Convergent parallel design used for the study

The quantitative and qualitative methods used for data collection are discussed in detail later in this chapter.

4.5.2 Why Mixed Methods Approach for the Present Study?

Before finalizing the design for the present study, the research problem was studied and considered carefully. The study proposed to find out if there exists a relationship between English teachers' CLAL and classroom assessment practices. One of the research objectives was also to trace the impact of the need-based TD programme in language assessment on teachers' performance in assessing language ability in the classroom. The impact could have been captured by using only qualitative or only quantitative means. Instead, the study used both kinds of methods. It was an informed and calculated decision to obtain more reliable data. For example, the CLAL survey would not have been enough to get information about teachers' assessment ability. Analysing teacher-made assessment tasks, their feedback on student performance, etc. qualitatively added a solid dimension to the authenticity, validity and reliability of the data. Similarly, the evaluation of assessments designed by the teacher could have been done alone qualitatively using a portfolio. But a quantitative grading of the assessments prepared before and after the intervention further substantiated the changes concretely.

4.6 Why Multiple Case Study Approach for the Present Study?

This project proposed to use an embedded multiple-case design because it suited the problem under study. As the study involved English teachers working in secondary schools, it was found appropriate and methodologically sound to study them in their natural setting. But since schools usually function under three different boards (CBSE, ICSE and State Boards) in the country, it was necessary to study teachers working in all the three types of schools. This could be accommodated by a multiple-case design. Moreover, it was felt that the multiple units and levels of analysis demanded by the study could only be met in an embedded case study. An embedded approach would gel well with the overarching Mixed Methods Approach adopted for the study.

4.6.1 Case Study

Qualitative in nature, case study is very difficult to define in specific terms. According to Yin (2003), case studies are favoured as a research strategy "when 'how' or 'why' questions are being posed, when the investigator has little control over events, and when the focus is on a contemporary phenomenon within some real-life context" (p. 1). In an applied linguistics context, a case study can be used for investigating a particular case which can be a teacher, a learner, a class, an institution or a group. In L2 situations, it is often used for tracking the language development, describing interventions, understanding and evaluating a particular case, etc. (McKay, 2006). In-depth analysis, context-specificity, involvement of individual characteristics and singularity are some of the key features of a case study (Duff, 2007). Apart from the specific nature (descriptive) of research questions, as pointed out by Johnson (1992, p. 91), a case must have a detailed description of the research context, the rationale for the selection of participants and a detailed account of their profile, the theoretical positioning of the study, the perceived role of the researcher, a thick description of the procedures of data collection, analysis and findings of the study, and a discussion on the validity and relevance of the study.

4.6.2 Multiple-Case Design

To overcome researcher and single informant biases involved in a single case study (Griffee, 2012), Yin (2003) suggests adoption of multiple-case designs. He talks about "analytic benefits" that multiple cases can offer. In other words, what he is hinting at is the quotient of generalizability. Though there are possibilities of having some differences

in the findings from individual cases, yet the common findings can easily lend themselves to greater claims for a larger population. Yin is right in asserting that 'replication' rather than 'sampling' is crucial to the construction of multiple-case designs. The logic here is that replications add to more reliable findings. However, the researcher may have to reconsider the number of variables to be included in the study. In single-case studies, an in-depth study of many variables is encouraged. However, a similar approach in the case of multiple-cases will give rise to feasibility issues. Yin also emphasizes that all the replications need not be 'literal replications'. A few cases may be literal and others can be theoretical. If a multiple-case design follows a proper replication pattern, the problem with the number of cases will not be an issue (Yin, 2003, p. 51). In addition, the researcher must decide whether the study will be holistic or embedded in nature. According to Scholz and Tietje (2001), "A holistic case study is shaped by a thoroughly qualitative approach that relies on narrative, phenomenological descriptions." (p. 9), and "Embedded case studies involve more than one unit, or object, of analysis and usually are not limited to qualitative analysis alone. The multiplicity of evidence is investigated at least partly in subunits, which focus on different salient aspects of the case." (pp. 9-10)

This study opted for an 'approximate' replication strategy. According to Abbuhl (2011), this kind of replication "involves repeating the original study exactly in most respects, but changing one of the non-major variables" (p. 298). In this case, the only variable that changed was the type of school (CBSE, ICSE and State Board) teachers belong to. The major variable, i. e., the TD programme on classroom assessment of language ability remained unchanged in all three cases. Moreover, since it was a mixed methods research, the quantitative and qualitative means of data collection were closely replicated.

Apart from what is mentioned above, a few other things were special about the procedure of replication adopted for this project. Rather than completing one study and then replicating it after a gap of weeks or months, replication happened almost simultaneously at three stages of data collection and ensuing analyses. The following diagram provides a clear idea about the process:
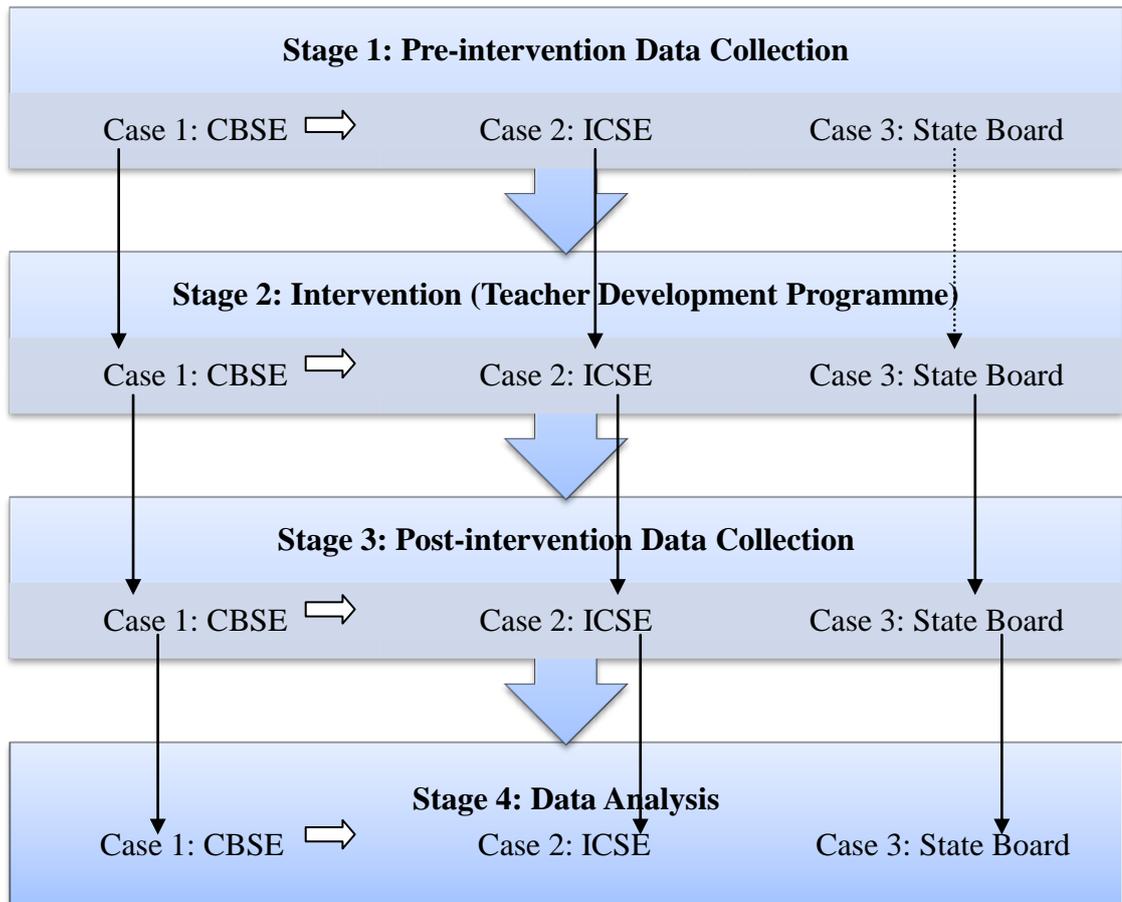
Figure 4.2: Process of data collection

4.7 Sampling for the Study

Since this study follows a mixed methods approach, sampling for the study also takes the same course. The plan of sampling is thoroughly grounded in several important considerations concerning the study. Hisse-Biber (2010, pp. 54 - 55) thinks that the following should top the list of such considerations in a mixed methods study:

- research questions
- adequacy of the sample size with regard to the main purpose of the study
- possibility of drawing conclusions in relation to the sampling design

The above-mentioned suggestions were placed in a framework, meant to be used for Mixed Methods research, called "Matrix Crossing Type of Sampling Scheme" (p. 284). Devised by Onwuegbuzie and Collins (2007), this is "a matrix that crosses type of sampling scheme (i.e., random vs. non-random) and research approach (qualitative vs.

quantitative)" (p. 284). The "Type 2" of the scheme involves random sampling for quantitative strand and non-random sampling for qualitative strand.

4.7.1 Sampling for the Survey

A simple random approach was followed to select the sample for the survey conducted during the first stage of the research. Since there was very little information available about the exact number of English teachers working in secondary schools in Andhra Pradesh, the total number of 'full-time' secondary school teachers, i. e. 198,812 (MHRD, 2009) was randomized to generate a list of 120 teachers across 23 districts of Andhra Pradesh. The following figure contains a part of the list:

Research Randomizer Results:
1 Set of 120 Unique Numbers Per Set
Range: From 1 to 198812 -- Sorted
Set 1

152		1
258		2
3460		3
3709	Adilabad	4
4385		1
5683		2
7491		3
7981		4
9268	Anantapur	1
10469		2
13336		3
13899		4
15041		1
16308	Chittoor	2
22031		3

Figure 4.3: List of teachers generated by the randomiser

As shown in the above figure, the list of random numbers, ranging from 152 to 22031 and mentioned on the left side, represents teachers. Since the number of teachers and thus, the range was fixed for each district, only the numbers, which were generated by the randomiser, were placed against the range for each district. So, the randomiser

generated only six teachers from Adilabad, which had the range of 6738 (1- 6738) in the list published by MHRD, six teachers from Anantapur, which had the range of 7621 (6738 – 14359), three teachers from Chittoor, which had the range of 10568 (14359 – 24927), and so on. The list of the 120 teachers, who participated in the AL survey, was generated in a similar manner for all the districts.

On the right side, there is another list of numbers, ranging from 1 to 4. Here, 1 and 2 represent state-run English medium and Telugu/Urdu medium school teachers, and 3 and 4 stand for CBSE and ICSE school teachers respectively. So, from Adilabad district, two teachers from state-run English medium, two from state-run Telugu medium, one each from a CBSE and an ICSE school participated in the survey. APPENDIX A contains the complete list.

The above-mentioned calculation was perhaps one of the very few options available for obtaining a list of teachers for the study because the list for the survey was generated from the total number of secondary school teachers, which included teachers from all disciplines. It was assumed that the proportion of the total number of teachers working in each district was equal to the proportion of the number of English teachers working in each district. Since there was no official information available about the exact number of English teachers working at secondary level in the state, no better feasible idea could be conceived to get a more accurate sample.

4.7.2 Case Selection for the Multiple-Case Study

A purposeful approach to sampling is adopted for selecting the three cases of the intervention study. Experts in case study research (Yin, 2003; Duff, 2008; Mills, Durepos & Wiebe, 2010) have discussed this kind of sampling. Since the nature of this multiple-case study is relational in nature and completely dependent on an intervention, it was decided to follow the nature of replication in the process of case selection. It has already been discussed in one of the previous sections of this chapter (4.6.2) that approximate replication, i. e., the replication design used for this study, emphasises selection of cases which are different from each other in terms of the variables the study plans to look into. Earlier, it was found from the state-wise survey of CLAL that secondary school English teachers working under three different boards, i. e., CBSE, state-board and ICSE, had differing average levels of CLAL. The assessment responsibilities and the school management system were also different. All these factors were considered during the

process. In addition, availability and willingness of teachers to participate in the study and readiness of the school management to allow the researcher to carry out the study for 18 hours were two crucial factors in the case selection. Finding teachers and making them agree to participate in the test-cum-survey was a difficult task. It was not possible to conduct the study in rural schools because of two reasons: amount of money and time required for travel and stay in the place, and getting access to secondary schools and willing teachers who would continue till the completion of the course. A detailed description about the selected schools and teachers is presented in the next chapter.

4.8 Description of Cases

Three school boards—CBSE, Andhra Pradesh Board of Secondary Education (APBSE) and ICSE—, represented respectively by three schools, comprised the three cases for the study. Further, two secondary level English teachers from each school were selected to participate in the study, which included a TD programme aiming to enhance their assessment ability. The assessment policies of the three boards along with assessment culture in all three schools and a brief profile of each teacher are presented in the following sub-sections.

4.8.1 Case 1

CBSE Assessment Policy

As shown in the table below this paragraph, an academic year is divided into two 'terms' and each term has two formative assessments (FAs) followed by a summative one (SA). A student got an overall grade based on their performance in all these six assessments. Each FA got 10% weighting and 30% weighting was given to each SA. So students' performance in classroom assessments (FAs) amounted to a total 40% weighting in the final calculation of grade, and these assessments are supposed to be designed, conducted, graded and evaluated by the teacher.

Table 4.1: Assessment responsibilities of CBSE teachers

Type of Assessment	Percentage of weighting in academic session	Month	Term wise weighting
	First Term		
FA – 1	10%	April-May	FA1 + FA2 = 20%
FA– 2	10%	July-August	
SA –1	30%	September	SA1 = 30%
	Second Term		
FA– 3	10%	October-November	FA3 + FA4 = 20%
FA– 4	10%	January- February	
SA– 2	30%	March	SA2=30%

According to the principles of CCE, which guided the assessments in the school, these FAs are expected to be integrated with classroom teaching. They should be diagnostic in nature and used for keeping track of learners' development of language proficiency. Teachers, after analyzing learners' performance, should offer constructive feedback to students. Teachers are also encouraged to employ a variety of assessment methods and assess oral language skills along with other communicative abilities.

The School and its Assessment Culture

A well-known co-ed CBSE school in Hyderabad formed the first case. It had around 1500 students studying in it. Most of these students were from middle and lower middle class families. The school had five working days per week. Though teachers working in this school were kept busy with classes, they were provided with some professional support by the school authorities. In fact, the principal of the school immediately agreed

to let two of his teachers participate in the TD programme offered by the researcher. He chose these teachers and informed the researcher that the selected teachers would be able to disseminate the knowledge gained through attending the programme among their colleagues. However, he asked the researcher to visit the school at a time convenient to both the teachers and conduct the sessions.

The principal wanted his teachers to update themselves professionally but he did not mention teachers' assessment abilities and FAs during the conversation with the researcher. The English teachers were quite occupied with teaching of English and an extra subject, along with other official duties. Most of them used paper-pencil tests and traditional questions on reading and writing for all the FAs. Even the prescribed oral assessment component was conducted through written tests. Except a participant-teacher, the others did not know much about FAs and how to conduct and use them. Grades mattered a lot to students and their parents. So teachers were asked not to be too strict while assigning grades to students. The school management did not take any special interest in providing all the English teachers with opportunities for professional development in assessment though general the school conducted some refresher courses on pedagogy in general. Only one teacher (*Teacher 1*) reported that she had attended a workshop conducted by a foreign trainer. They did not have time to participate in teacher development programmes offered by ELT@I Hyderabad Chapter, British Council and RELO.

Teacher 1

This teacher was a 42 year old female with 15 years' experience as an English teacher. She possessed an M. A., an M. Sc. and a B. Ed. degree. She had been working in that school for about eight years. She could read, write and speak quite fluently while teaching in the classroom and also during the intervention. She was enthusiastic about the intervention programme and wanted to know more about the programme before the start. She was the most experienced English teacher in the school and taught in classes IX and X. She had undergone some training programmes in language assessment and was aware of CCE and the requirements related to classroom assessment. She wanted to know more about designing appropriate assessment tasks, integrating assessment with teaching and providing feedback on students' assessment performance.

Teacher 2

A 33 year old male, the second teacher had an M. A. and a B. Ed. degree and six years of experience in teaching English. He had been in that school for six years and was teaching classes VIII and IX. He informed the researcher that when he had joined the school, he had a lot of difficulties with using the English language. He had been working on his English since then. He faced some difficulties in terms of using accurate English while interacting with the researcher and also had some problems in writing. He agreed to join the TD programme because he thought it might help him become a better teacher. Also, he wanted to know from the researcher how he could improve his English. Apart from teaching English, he taught science to students of class VIII. He had very little idea about CCE and classroom assessment. For classroom assessment purposes, he used questions from question banks sold in the market.

4.8.2 Case 2

APBSE Assessment Policy

The AP Board followed an adapted version of CCE as its assessment policy for secondary level classes. The English teachers had to take care of the formative assessments for the classes they were teaching. As mentioned below, the prescribed assessments were divided into FAs and SAs. There were four FAs and three SAs conducted during an academic year where the FAs accounted for 20% of the total marks. The Board very clearly directed school principals and teachers to integrate FAs with classroom teaching and make use of assessment methods like observation, project work, written portfolio assessment, etc. along with slip tests. It also emphasized the use of rubrics for FAs. Teachers had to maintain a record of students' performance in FAs and assign marks to them accordingly.

Table 4.2: Assessment responsibilities of State Board (Andhra Pradesh) teachers

Academic Standards		FORMATIVE ASSESSMENT						SUMMATIVE ASSESSMENT			
		Weight-age	Obser-vation	Note books	Pro-jects	Slip Test	Total	Weight-age	Oral	Written	Total
Listening & Speaking		10%	5				5	10%	10		10
Reading Comprehension		10%	5				5	15%		15	15
Conventions of Writing		10%		5			5	5%		5	5
Vocabulary		10%				5	5	10%		10	10
Grammar		10%				5	5	10%		10	10
Creative Expression	(a) Oral							10%	10		10
	(b) Written	30%		5		10	15	40%		40	40
Projects		20%			10		10				
Total		100%	10	10	10	20	50	100%	20	80	100

(Obtained from http://bseap.org/)

The School and its Assessment Culture

The state-board government-run English-medium school is in Hyderabad (urban). Around 600 children study in this school. Most of these children belong to lower middle class families. Of late, the school has not been getting much aid, in terms of financial support, from the government. The number of students studying in that school has been on the decline. However, all the teachers working in this school were formally trained and experienced. The principal was quite active and wanted his teachers to participate in professional development activities. So he immediately agreed to the proposal of the researcher and asked two teachers to participate in the intervention programme. Much like the principal of the aforementioned CBSE school, he asked the researcher to schedule his sessions when the two teachers did not have classwork.

The school offered more flexibility than the aforementioned CBSE school to teachers in terms of conducting FAs. In addition, a senior teacher from the school was also a part of the district committee of question paper designers. Though he was not an English teacher, he encouraged his colleagues to develop their assessment ability. The English teachers were very unhappy with the new assessment policy (CCE). They believed that the policy burdened them with more assessment responsibilities because they did not know how to assess students' progress without spending too much time on it. The

teachers requested the researcher to help them find ways so that they could carry out FAs in the classroom without much struggle. The principal of the school also agreed with his teachers and informed the researcher that his teachers needed some orientation in FA.

Teacher 3

This teacher was 52 years old and had more than 20 years of teaching experience. Her educational qualifications included an M. A. and a B. Ed. degree. She had been teaching in that school after getting into the field of teaching. One of the senior teachers in the school, she was proficient in using English. She was easy-going and informed the researcher in advance that she wanted to improve her teaching and assessment practices and would like to learn new things about teaching and assessment. She was unhappy about the new assessment policy and the lack of training in assessment for teachers in government-run schools. She taught students of classes IX and X students but did not know how to deal with new assessment policy.

Teacher 4

The second teacher was 42 years old and had been teaching English for 10 years. He had completed M. A., M. Phil. and B. Ed. before getting into teaching. He was teaching English to students of classes VIII and IX. He informed the researcher that he had done B. Ed. without attending any class and thus, he had learnt very little about teaching English during his pre-service training. Whatever teaching skills he had, he thought he had acquired through his experience as a teacher in that school. He had some problems with communicating in English. He was aware of that and requested the researcher to provide him with some materials so that he could improve his English. Though he was not really interested in participating in the intervention programme and he was asked to be a part of the programme by his principal, soon he started taking interest in the programme and worked seriously on the tasks during the training period. He requested the researcher to meet him outside the school and sought his advice regarding ways to improve his spoken English.

4.8.3 Case 3

ICSE Assessment Policy

The assessment system in ICSE schools is different from that prescribed by the CBSE. Classroom assessment of language skills got 10% weighting in English. Considering that the subject was of 200 marks, the percentage of marks assigned to FAs was quite low. Even the 20 marks assigned for internal assessment were confined to the assessment of listening and speaking skills in a prescribed manner. The following table provides the details:

Table 4.3: Assessment responsibilities of ICSE teachers

Subject Area	External Assessment	Internal Assessment	
English Language (100 marks)	80 marks	20 marks	Thrice a year for class IX (Total 20 marks) Twice a year for class X (Total 20 marks)
		Listening Skills: note-making, general comprehension check (written) Speaking Skills: presentation, interview	
English Literature (100 marks)	80 marks	20 marks	

So the secondary level English teachers in the school were expected to select a passage, read it aloud twice, ask students to make notes, design a few comprehension questions to assess students' listening comprehension skills, ask students to make an oral presentation, evaluate it and interview each student for about 3 minutes on the topic presented by the student.

The School and its Assessment Culture

The school was one of the most famous ICSE schools in Hyderabad. Around 1500 students were studying in that school. The students belonged to different classes of the society. The school waived the tuition fee for students from poor families. Most of the teachers who worked there had B. Ed. In the school, English as a subject got a lot of

importance. Writing and speaking correct English got utmost importance. The school management kept all the teachers very busy with academic and co-curricular activities. The principal was not willing to allow her teachers to participate in the intervention programme offered by the researcher. She thought any such programme would disturb the functioning of the school.

Assessment of language skills was confined to traditional tests in the school. As the syllabus contained a lot of literature, language was never the first priority of teachers. The tests were loaded with writing and grammar. The teachers felt that CCE was a total waste of time and that it would promote unhealthy practices. They wanted the government not to impose the framework used in CBSE schools on their school. Tolerance of error was unacceptable to most English teachers. Even the principal was unhappy with the growing emphasis on 'communication skills'. She thought the English language was losing its purity. She thought that her teachers were quite capable of preparing their students for the examination and was sure that the English teachers working in her school did not need any training in language assessment. She was of the opinion that any such training would not improve their ability to help students score better in the examination.

Teacher 5

The first teacher in this school was a 65 year old lady with 35 years of experience in teaching English. She possessed an M. A. and a PGDTE (CIEFL, Hyderabad) degree. She taught at various levels including primary, secondary and college. She had been in that school for about 10 years and was teaching classes IX and X. She had a strong belief system about language teaching and learning. She was very proficient in English and expected every student to be accurate while speaking and writing. During a discussion before the commencement of the intervention programme, she informed the researcher that she was very unhappy with the change in approaches to teaching and assessment. She told the researcher about her belief in grammar-based teaching, her doubts about the new grading policy proposed by CCE and convictions about an examination-oriented teaching. She was sure that the intervention programme in assessment would offer very little for her to learn.

Teacher 6

The second teacher from the ICSE school was a 37 year old lady with 12 years of teaching experience. She was an English graduate (B. A.) and had pre-service training (B.

Ed.) before joining as a teacher. She handled classes VIII, IX and X. She was proficient in English and followed *Teacher 5* as a guru. She informed the researcher that *Teacher 5* was considered to be an ideal English teacher also by all other English teachers in the school. Like *Teacher 5,* she also believed strongly in accuracy and wanted all her students to use correct English in writing and speaking. Though she had virtually no training in language assessment, she thought she did not require any training in language assessment. She told the researcher directly that she had agreed to be a part of the intervention programme because the school management wanted her to do so. She was dependent on sample papers for preparing her students for final examinations and was convinced that same approach could be adopted for classroom assessments.

4.9 Methods of Data Collection

After meticulously analysing the research questions, the theoretical framework and the research methods used by previous researchers in the area, certain research methods were found apt and thus, were chosen for the present study. They included a CLAL survey instrument, a teacher belief questionnaire, a rating scale, interviews, field notes, reflective journals and an electronic portfolio. These methods are described in detail in the following sections of this chapter.

4.9.1 CLAL Survey Instrument

The instrument used for the survey was called *CLAL Survey Instrument.* The decision was made after reviewing the research in the concerned area. Whereas Dornyei and Taguchi (2010) have problems with calling the survey instrument a questionnaire if it tests "the respondent's competence in performing certain tasks" (p. 5), Brown (2001) and Mackey and Gass (2005) do not mention anything against it. Mackey and Gass (2005) in fact talk about specialized types of questionnaires used for specific research purposes. Moreover, the most famous survey instrument ("Teacher Assessment Literacy Questionnaire") designed by Plake and Impara (1992) for assessing teachers' assessment literacy is called a questionnaire. A similar kind of instrument, developed by Mertler (2002), is called "Classroom Assessment Literacy Inventory".

4.9.1.1 Rationale for the Use of the Survey Instrument

There were good reasons behind the use of the survey instrument and for not using existing ones from the literature. Firstly, since the study was based on a claim about the

classroom assessment literacy of secondary school English teachers, it was necessary to empirically support any such claim. The state-wise survey provided the required data. Secondly, the existing classroom assessment inventories (such as those by Plake and Impara (1992), Mertler (2002), etc.) used a lot of technical terms and the questions included in them did not suit the context in which the current study was based. Considering the kind of assessment training teachers got and the kind and amount of assessment practices they were required to engage in, it would have been unfair to assess the CLAL using those instruments. Even B. Ed. (English) programmes did not impart adequate training in classroom assessment of the English language. Moreover, in schools, teachers were expected to assess their students as a part of the formative or internal assessment, and they may not require extensive theoretical knowledge in aspects of language assessment to carry out those assessments.

4.9.1.2 Construction and Piloting

The survey instrument was designed keeping in mind the knowledge, skills and abilities secondary school English teachers must possess in order to carry out classroom assessments. The instrument went through a long and tedious process of editing, reviewing, re-drafting and piloting before taking its final form. It took almost two months to get the instrument ready for piloting. The seven-page long survey instrument had seven sections numbered numerically from 1 to 7. Some sections were further divided into sub-sections.

In the first draft of the instrument (APPENDIX B), there were 19 questions in the instrument and most of the questions were MCQs and a few were True-False type. Technical terms in assessment like 'validity', 'reliability', 'authenticity', 'washback' and 'practicality' were used in almost all the questions. The aim was to obtain teachers' ability to understand and apply these aforesaid principles in classroom assessment contexts. But the researcher, while conducting a workshop for a group of 18 secondary school English teachers in another context, realised that teachers were not aware of the meaning of these terms. Thus, in the second draft, all these terms were removed from the instrument. In the next draft of the instrument, there were no technical terms and all the questions were based on small assessment tasks or situations, quite similar to those used by secondary school English teachers during classroom assessments. Only the last question, i. e., the seventh one, required teachers to know the steps to carry out classroom assessments effectively. There was no assessment context for that question. In

addition to the above-mentioned changes, the questions with more than one answer were changed into right-wrong type questions with just one right answer for each question.

The instrument was piloted with 20 target group teachers working in CBSE, state-board and ICSE schools across Andhra Pradesh. Out of 20, 10 teachers were from state-board schools (both English and Telugu medium) and five each from CBSE and ICSE schools. The researcher visited all these schools and was present with the teachers when they responded to the questions in the instrument. He noted down the problems faced by the teachers and marked the areas on the instrument which might need some reviewing.

4.9.1.3 Final Draft

Just one change was made in the piloted CLAL survey instrument. In the fifth question, it was observed during piloting that some teachers took options to the question as sections of the question and responded to them. There was a little bit of confusion. To improve clarity in the instruction, one sentence— "Choose the best option out of 'a', 'b', 'c' and 'd'."— was added to the instruction for the question. (See APPENDIX C for the final draft.)

The reliability of the instrument was estimated using KR21 formula. As shown below, at first, standard deviation and mean were calculated using an online calculator and later an MS Office Excel spread sheet (obtained from the website http://languagetesting.info) was used to calculate the KR21 for the piloted instrument which turned out to be .73, an acceptable score for a small-scale survey containing 38 items. The calculations are presented below:

Standard Deviation Calculator

To Calculate Mean, Variance, Standard deviation :
Enter all the numbers separated by comma ','.
E.g: 13,23,12,44,55

12,13,14,12,18,17,19,17,19,13,21,22,19,20,
24,28,28,29,30,20

[calculate]

Results:

Total Numbers:	20
Mean (Average):	19.75
Standard deviation:	5.69279

R2C3 f_x =(1-((mean*(k-mean))/((SD*SD)*k)))*multiplier
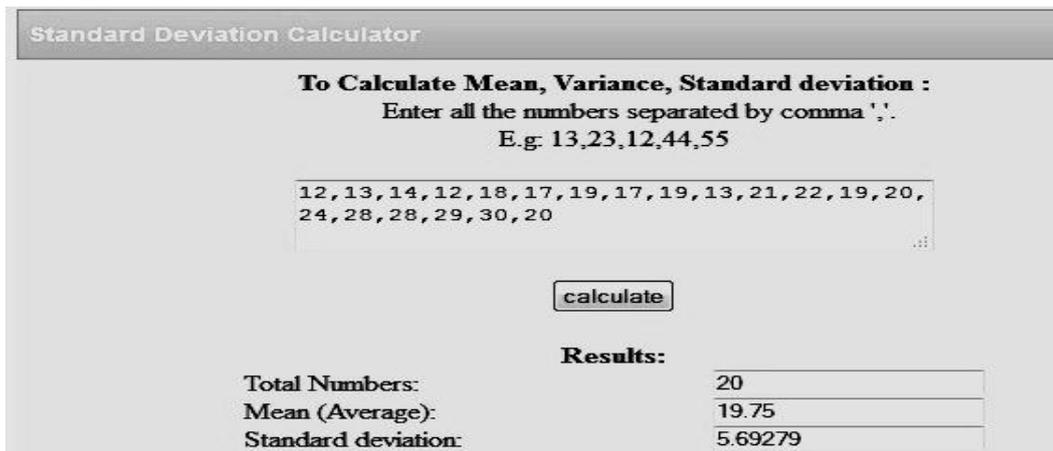
	1	2	3	4	5	6	7
1	k here ->	38					
2		37	0.73	<-- KR 21			
3		1.027027027					
4	SD here ->	5.69279					
5	mean here ->	19.75					

Figure 4.4: Calculation of reliability for the survey instrument

4.9.2 Teacher Belief Questionnaire

A teacher belief questionnaire (see APPENDIX E) was developed by the researcher to trace the beliefs of the six teachers (who participated in the intervention) before and after the TD programme. It contained 27 items and 23 of those were meant to elicit the teachers' beliefs about classroom assessment of language ability. The first four items were used to obtain information about their experience and perception of training in assessment during B. Ed. and in-service days. The rest of the items were developed under three broad categories: Classroom Assessment, Principles of Assessment and Teachers' Assessment Ability. It must be mentioned that these categories were never meant to be water-tight compartments. Rather, the intention was to ensure smooth and meaningful processing and analysis of the acquired information under these categories. However, this categorisation was not followed in the arrangement of the questionnaire administered to the teachers. In the questionnaire questions 9, 18, 19, 25 and 26 belonged to Classroom Assessment; 7, 8, 10, 11, 14 and 16 could be placed under Principles of Assessment; and 5, 6, 12, 13, 15, 17, 20, 21, 22, 23, 24 and 27 comprised the category Teachers' Assessment Ability.

The questionnaire was scored on a five-point Likert scale comprising *SA* (strongly agree), *A* (agree), *D* (disagree), *SD* (strongly disagree) and *DN* (do not know). Mean values were assigned to these options and the responses were coded and the analysis is presented in the next chapter. The teachers had to choose one of these options to indicate their belief about the statements in the questionnaire. The questionnaire was used for the study after it was scrutinised by three university teachers who had experience in training Masters and Research students at university level.

4.9.3 Semi-Structured Interviews

Interview was one of the important tools of data collection used in this study. The interviews were semi-structured in nature. In this kind of interview, as rightly explained by Dornyei (2007), "[t]he interviewer provides guidance and direction (hence the '-structured' part in the name), but is also keen to follow up interesting developments and to let the interviewee elaborate on certain issues (hence the 'semi-' part)". The six teachers, who participated in the intervention programme, were interviewed before, during and after the programme. Three different sets of questions were used for eliciting the required information. The first set was used as a follow-up procedure after they had filled out the CLAL survey instrument and the teacher belief questionnaire. The aim was to collect more information about the teachers' knowledge and belief about classroom assessment of language ability. The teachers' views and experiences during the intervention programme were recorded using the next set of questions. The last set was employed to get information, once again, about the teachers' knowledge and belief about classroom assessment of language ability after the intervention was over. (See APPENDIX M for the interview questions.)

4.9.4 Rating Scale

A rating scale (see APPENDIX F) was designed specifically keeping in mind the objectives of classroom assessment proposed by CCE. A set of ten important characteristics of classroom assessment tasks was presented and four levels were created for each characteristic of the task. These levels were, in fact, points on a continuum with '1' representing the least fulfilment and '5', the best fulfilment of the characteristic or condition. This rating scale was used to find out the quality of assessment tasks constructed by the teachers before and after the intervention programme. Before using it for the aforesaid purpose, the scale was evaluated by three university teachers who had

been teaching language assessment. Their feedback helped to sharpen the focus of the instrument and improve its content.

4.9.5 Field Notes

Tavakoli (2012) asserts that field notes are "the most common method of recording the data collected during OBSERVATION" (p. 228). In this study, field notes were used to note down relevant details about teachers' classroom assessment ability while observing classes of six teachers before and after the intervention. The researcher did not make descriptive notes nor any deliberate attempt to interpret the observation data separately. He was a non-participant observer and made "field jottings" (Bernard, 2000, p. 357) using a note pad. The collected data were not the main source of data in the study. The researcher made use of field notes (see APPENDIX L) to corroborate findings about the teachers' level of assessment literacy obtained through the CLAL survey instrument, teacher beliefs questionnaire and assessment tasks designed by the teachers.

4.9.6 Reflective Journals

The six teachers, who participated in the intervention study, were requested to maintain a reflective journal during the training period. Journals are considered to be one of the rich sources of data in qualitative research. Richards and Farrell (2005) hold that keeping a journal helps teachers in professional development. In the current study, the reflective journals (see APPENDIX I) provided necessary data about teachers' personal experiences and their views about the TD programme they were undergoing. The researcher used the data to give direction to the programme and trace their progress. Simultaneously, journal-keeping was expected to help teachers to reflect on what and how they were learning the process and nature of assessment during the programme.

The teachers were given clear directions in what and how to write their journals. They were asked to reflect on the following questions at the end of each week of training and write a paragraph or two on their response:

- What did you find useful and new about this week's training?
- How did you find the tasks used last week during the sessions? Were they interesting/ difficult/ useful?
- Do you want to suggest anything to improve the programme?

4.9.7 Electronic Portfolio

Electronic portfolio, also known as e-portfolio or digital portfolio, is a collection of evidence stored in electronic form. It is generally a web-based method of storing data but can also include offline storage of data. In the present study, an electronic portfolio was maintained by the researcher. He stored scanned copies of assessment tasks prepared by the teachers before and after the intervention programme and real classroom assessment tasks and a voice-recorded reflective account of his own experience during the programme. In addition, copies of weekly reflective journals submitted by the teachers and audio files containing interviews with them before and after the intervention were also included in the collection.

It was found that the electronic portfolio facilitated easy and systematic storage of important data. It was also easy to access every single piece of evidence which, in turn, helped to avoid any kind of confusion during the analysis. Separate folders were created to store data obtained from each teacher.

4.10 Procedure of Data Collection

The process of data collection lasted for around five months and had five stages. These stages covered the data collection before, during and after the intervention. This pre-planned division into stages was meant to keep the process confusion-free and made the presentation of the process in the dissertation easy. The following sections in this chapter contain detailed description of the process.

4.10.1 First Stage: CLAL Survey

In the first stage, the CLAL survey instrument was administered to a group of 120 in-service secondary school English teachers from Andhra Pradesh. The aim was to get some idea about the average CLAL of teachers working in Andhra Pradesh. The survey was necessary to support the basic premise on which this research was based.

The researcher collected the postal address of one person he knew in each district (except Hyderabad) of Andhra Pradesh. These twenty-two persons were sent the required number (the number of teachers selected as sample) of copies of the instrument by post. They were asked to immediately contact the researcher in case teachers faced any difficulty while responding to the instrument. After collecting the names and phone numbers of the teachers and schools they had to visit and taking prior appointment with

these teachers, they visited the teachers and got the instrument filled out by the teachers. Only a few teachers from Anantapur, Adilabad and Khammam district reported some confusion while responding to the questions. They wanted to know whether they were required to respond to the questions in the tasks. The researcher spoke to them on phone and helped them with the instrument. Though teachers from other districts did not report any problem in completing the survey, later it was found that quite a few teachers completed the tasks on which the survey questions were based. Some even went on to write an essay. However, all the teachers responded to all the questions in the instrument. This happened because the people, who were responsible for meeting the teachers and obtaining their responses, were not much aware of the content of the questionnaire. But this was a practical difficulty that could not have been avoided.

It took almost 20 days to collect the responses of all the teachers from 22 districts. The researcher did not take help of anyone for collecting data from 13 teachers in Hyderabad. He visited all these teachers at their respective schools and was present with the teachers when they responded to the survey instrument. He made sure that teachers had no problem in understanding the questions and what was expected of them. Most of the teachers took about 20-30 minutes to fill out the survey instrument. A report of the survey is presented in the next chapter.

In addition to the survey, the B. Ed. syllabi of three B. Ed. programmes in the English and Foreign Languages University (EFLU), Osmania University (OU) and Regional Institute of Education, South India (RIESI) were evaluated. The aim was to gather information about the focus on language assessment in these pre-service teacher education syllabi. The analysis is presented in the next chapter.

4.10.2 Second Stage: Classroom Observation and Assessment Task Analysis

In the second stage, six teachers, i. e., two each from a CBSE, a state-board and an ICSE school, were selected and treated as three different cases for the study. Three classes each of all the six teachers were observed. The intention was to collect information about the following aspects:

- the kind of assessment methods and techniques the teacher was using
- whether the methods were integrated with teaching
- the manner of providing feedback

Rationale for classroom observation

Since FA is about integration of classroom teaching and assessment, information about a teacher's approaches to CA can be deduced from the way they handle classroom teaching. Therefore, even those days during which teachers do not assess become relevant for observation. Also, a lot of assessment techniques for formative purposes like asking questions, providing students with opportunities to use language, making use of appropriate context while presenting language items, offering feedback, etc. are commonly used in classroom teaching situations. The aim was to capture the above mentioned details through classroom observation. In addition, it was an opportunity to see how teachers' knowledge and beliefs were getting reflected in their practice.

Though the earlier plan was to video-record and analyse the classes, the researcher took only field jottings while observing the classes. It happened because the management body and the teachers of one school opposed the idea of video-recording the classes as it was a girls' school. They had the fear that parents might object if they permitted the researcher, who is a male, to record the classes in which there were adolescent girls. Since it was necessary to maintain uniformity in data collection across all three schools, it was decided to cancel the video-recording of classes in the other schools. Only a single class of a teacher was observed on one day to avoid the impact of any affective factors like fatigue and anxiety on the researcher. It took almost eleven days to complete the observations.

Apart from classroom observation, assessments designed by the teachers and information about the teachers' assessment duties and responsibilities were collected and analysed. At first, the plan was to collect real classroom assessment tasks used by the teachers for their classroom purposes and analyse them. But when the teachers were requested to give copies of some of these tasks to the researcher, three out of six teachers politely declined. They had many excuses which forced the researcher to confine his options to collection of assessment tasks readily designed by the teachers in a formal setting. These assessment tasks were rated using a pre-determined rating scale. The details about the rating scale have been already discussed in one of the earlier sections (4.8.4) of this chapter. Two Ph. D. research scholars (who had training in language assessment) and the researcher rated each task using the scale. Then a quantitative analysis was carried out. In addition to the above-mentioned analysis, the real classroom assessment tasks, which were collected from three of the six teachers, were qualitatively

analysed. Inferences were drawn from both types of analyses and used in the fourth stage for developing the TD programme on classroom assessment and later, for tracing teacher change.

As mentioned in the first part of the above paragraph, information about teachers' assessment duties and responsibilities was collected from assessment-related documents from the schools and websites of CBSE, Board of Secondary Education of Andhra Pradesh (BSEAP), ICSE, NCERT and MHRD. The information was fed into the TD programme.

4.10.3 Third Stage: Data about Teachers' CLAL and Beliefs

In this stage, data were collected from six teachers, who participated in the TD programme later, about their CLAL and beliefs about classroom assessment. These teachers were not part of the state-wide CLAL survey. The CLAL survey instrument, already used for the state-wise survey, was again employed to acquire information about the participating teachers' CLAL levels. The teacher belief questionnaire was used to obtain information about their beliefs. Both the survey instrument and the questionnaire were administered to each pair of teachers- one pair each from a CBSE, a State Board and an ICSE school- separately. At first, each pair completed the survey and then filled out the belief questionnaire. They were allowed to discuss with each other only the instructions given in the survey instrument for each question.

Each teacher was interviewed immediately after completing the beliefs questionnaire. The interviews were audio-recorded and later, selectively transcribed and utilised during the last stage of analysis in which teacher change was traced. Each interview lasted for about 20 minutes and during the sessions, some of the teachers went off the track while responding to the researcher's questions. The researcher never directly stopped them from deviating. He made use of some instant questions to get them back to the topic of the researcher's interest.

4.10.4 Fourth Stage: Intervention

In the fourth stage, an 18-hour long language assessment literacy development programme was developed on the basis of the information gathered from the analysis of data collected through the state-wise survey, the analysis of the B. Ed. syllabi and the assessment responsibilities of the teachers. This programme, designed for the selected

teachers, was transacted separately with each pair of teachers. All the three pairs were trained for eighteen hours and at an average of three hours per week. Care was taken to maintain uniformity in all aspects of training across the three pairs.

During the training period, the teachers were asked to maintain a reflective journal. The researcher collected the journal entry of each teacher at the end of each week's training. All the teachers wrote less than a typed-page as part of their weekly entry. The researcher scanned each entry and added all the entries to his electronic portfolio. Also, the researcher, after each session of training, voice-recorded his experience during the session and added the same to the electronic portfolio. The teachers were given oral feedback on their task performance during each session. In addition, the teachers were engaged in self- and peer-assessment of appropriateness and quality of the tasks. After each training session, there were teacher-initiated informal discussions on problems and issues related to classroom assessment practices in their respective schools. More information about the intervention is provided in the next chapter of the dissertation.

4.10.5 Post-Intervention

Since the research aimed to explore the relationship between CLAL of teachers and their assessment practices and it involved comparison between the teachers' assessment literacy and classroom assessment practices before and after the intervention, the collection of data before and after the intervention had to follow a similar course. So the CLAL survey instrument and teacher belief questionnaire were administered again to all the six teachers. Once again, they were asked to design classroom assessment tasks and these tasks were evaluated by the same two Ph. D. scholars and the researcher separately with the help of the same rating scale. A few real tasks were also analysed qualitatively. The teachers were interviewed immediately after completing the belief questionnaire and the set of questions, which were used before the intervention, were used for the purpose. The researcher then observed three classes of each teacher and made field jottings about their classroom assessment practices.

4.11 Conclusion

This chapter has discussed the research design and methods and procedure of data collection. It presents the methodological foundation on which the study was developed. In the next chapter data about the state-wise survey of CLAL, the intervention and the teachers' CLAL are analysed and interpreted.

CHAPTER 5

IMPACT OF THE TEACHER DEVELOPMENT PROGRAMME ON TEACHERS' CLASSROOM LANGUAGE ASSESSMENT LITERACY

5.1 Introduction

This chapter presents the data about the Teacher Development (TD) programme and its impact on CLAL of the teachers. In the first part of the chapter, reports of an evaluative study of the assessment components in three B. Ed. programmes and a CLAL survey of secondary school teachers in Andhra Pradesh are presented. This part is followed by a detailed account of the intervention programme. The last part of the chapter presents an account of the pre- and post-intervention CLAL levels along with the changes observed in all the three cases.

5.2 Language Assessment Components in B. Ed. Programmes: An Overview

To add validity to the claim that secondary school English teachers in India do not get adequate training in assessment, the syllabi of four well-known B. Ed. programmes were analyzed. The aim was also to support the TD programme on classroom assessment of language ability designed for the intervention study which forms the core of this research project. Though the TD programme focused on the needs of a very small group of teachers from Andhra Pradesh, the results of this analysis showed that such programmes should be developed and offered to secondary school English teachers working in different types of schools.

The B. Ed. syllabi of the English and Foreign Languages University (EFLU), Hyderabad, Osmania University, Hyderabad and Regional Institute of Education (RIE), South India were analysed and evaluated to find out how much importance was given to language assessment training in these programmes. Since these B. Ed. programmes are quite well-known and popular among aspiring English teachers in Andhra Pradesh, they were included in the analysis. The following table contains the criteria of evaluation and shows how each programme fares against the criteria.

Table 5.1: Analysis of assessment components in B. Ed. syllabi

Description	EFLU	OU	RIE
Does Language Assessment/Evaluation form a part of the syllabus?	Yes	Yes	Yes
Are the basic principles of language assessment included?	Yes, but given very little importance	No	No
Is there any focus on classroom assessment?	No	No	Yes
Is there any component on language proficiency?	No but "proficiency test"- a component	No	No but "assessing proficiency"- a component
Is there any focus on assessment criteria?	Yes, but limited to only Speaking and Writing	Yes, but very general in nature	Yes, but limited to "scoring criteria"
Is there any component on different methods of assessment?	No	Yes, but referred to as "tools"	Yes, but nothing explicit
Does assessment of individual language skills and components feature in the list?	Only Speaking and Writing	No	Only skills but not components
Is there any focus on providing feedback to students about their performance?	No	No	Yes
Are analysis and interpretation of assessment results included?	Only statistical analysis	Yes	No
Does the syllabus have anything on evaluating and improving assessments?	No	No	No
Weighting (in percentage) given to language assessment in the total syllabus	Around 1.7%	Around 0.83%	Around 1.6%

The table above gives some idea about how assessment training is provided to English teachers in B. Ed. programmes. All the three syllabi pay very little attention to language assessment and even less to classroom assessment. Important aspects of language

assessment like basic principles of assessment, concept of language proficiency, methods of assessment, assessment of individual language skills and components and providing feedback on student performance feature sketchily in the syllabi. None of the syllabi contains anything on evaluation and improvement of assessment tasks. To add to the worries is the weighting given to language assessment in the whole B. Ed. syllabus. The syllabi used in B. Ed. colleges across Andhra Pradesh and elsewhere may not possibly exhibit a very different trend. It may be impressionistic but training in language assessment remains ignored in most B. Ed. curricula across the country. Though all these programmes have components on "educational evaluation" in their syllabi, they cannot be expected to help trainees much with assessment and evaluation of language ability. A CLAL survey covering Andhra Pradesh was conducted by the researcher. The survey results, which are discussed later in this chapter, showed that most secondary school English teachers possessed inadequate levels of CLAL. Thus, the present study stands justified.

5.3 CLAL Survey: A Report

A state-wise CLAL survey was conducted to assess the average level of CLAL of in-service English language teachers across the state of Andhra Pradesh. Details about the participants and the instrument that was used for collecting data have been already discussed in Chapter 4. The results of the survey were intended to support a primary assumption about the CLAL of English language teachers across the state. Since this survey only supported the main study and was not in itself the main study, only a brief report of the survey is presented below.

To make the analysis of data convenient and obtain a clear picture about the CLAL of teachers, three levels- "Appropriate", "Average" and "Limited"- were created and each level was described as presented below:

Table 5.2: Description of CLAL levels

Level	Descriptor
Appropriate (80%-100%)	• Has a basic understanding of the basic principles of language assessment required for constructing classroom assessment tasks • Knows how to plan, carry out and evaluate a classroom assessment • Shows an understanding of how to interpret assessment results and offer feedback to students on their performance
Average (60%-80%)	• Has some idea about the basic principles of language assessment required for constructing classroom assessment tasks • Shows some understanding of how to plan, carry out and evaluate a classroom assessment • Knows a little about interpreting assessment results and offering feedback to students on their performance
Limited (Up to 60%)	• Has a very limited understanding of the basic principles of language assessment required for constructing classroom assessment tasks • Knows very little about how to plan, carry out and evaluate a classroom assessment • Shows very limited understanding of how to interpret assessment results and offer feedback to students on their performance

As mentioned in the above table the level 'Limited' covered scores below 60% of the overall score, i.e., score below 23 (as the total score is 38). The next level comprised scores between 60% and 80%, i.e., scores between 24 and 30. The last level, i. e., 'Adequate', included scores between 80% and 100% which were equal to 31 and 38 respectively. There were several reasons behind the above division into levels and the development of the corresponding descriptors:

- It was decided during the creation of the survey instruments that a teacher who would be able to complete the survey with very few wrong answers should be considered to have adequate level of CLAL.
- The implication of the above statement is that those who fall below the adequate level should be considered as average or below average.
- The survey instrument did not use any technical terms and was based exclusively on the assessment duties expected of the target group of teachers.

- The items in the survey did not try to test anything beyond what teachers would need to know for carrying out assessments effectively.

5.3.1 Analysis of Survey Results

The scores of all the participating teachers on the CLAL survey instrument were statistically analysed using descriptive statistics. The mean scores of the CBSE, ICSE and State Board teachers were found to be 22.33, 20.26 and 20.81 respectively. The overall mean for the total 120 teachers was 21.05 (see APPENDIX D for all the scores). The three mean scores for teachers belonging to the three different boards were statistically compared individually with the other two using t-tests.

Table 5.3: Inter-group comparison of mean scores

Groups	p-value at p <0.05 (one-tailed)
CBSE-ICSE	0.079905
State Board-CBSE	0.127566
ICSE-State Board	0.340198

(The p-values were calculated using the online calculator available at http://www.socscistatistics.com/tests/studenttest/Default2.aspx)

The p-values for the above-mentioned three pairs— 'CBSE-ICSE', 'State Board-CBSE' and 'ICSE-State Board'— were found to be 0.079905, 0.127566 and 0.340198 respectively at $p<0.05$. The calculation followed a one-tailed hypothesis since the claim was that English teachers working in schools— irrespective of the boards under which they worked— had an inadequate level of CLAL, and this hypothesis was directional. The above-mentioned *p-values* proved that the differences between the average CLAL levels of teachers working under three boards were found to be statistically insignificant. In other words, teachers' abilities to carry out classroom assessments were more or less equal, irrespective of the boards under which they worked.

As indicated in the first part of the section, the scores of the teacher-participants were categorised as per the levels created earlier in the section. The following graph presents the categorisation of scores into levels:
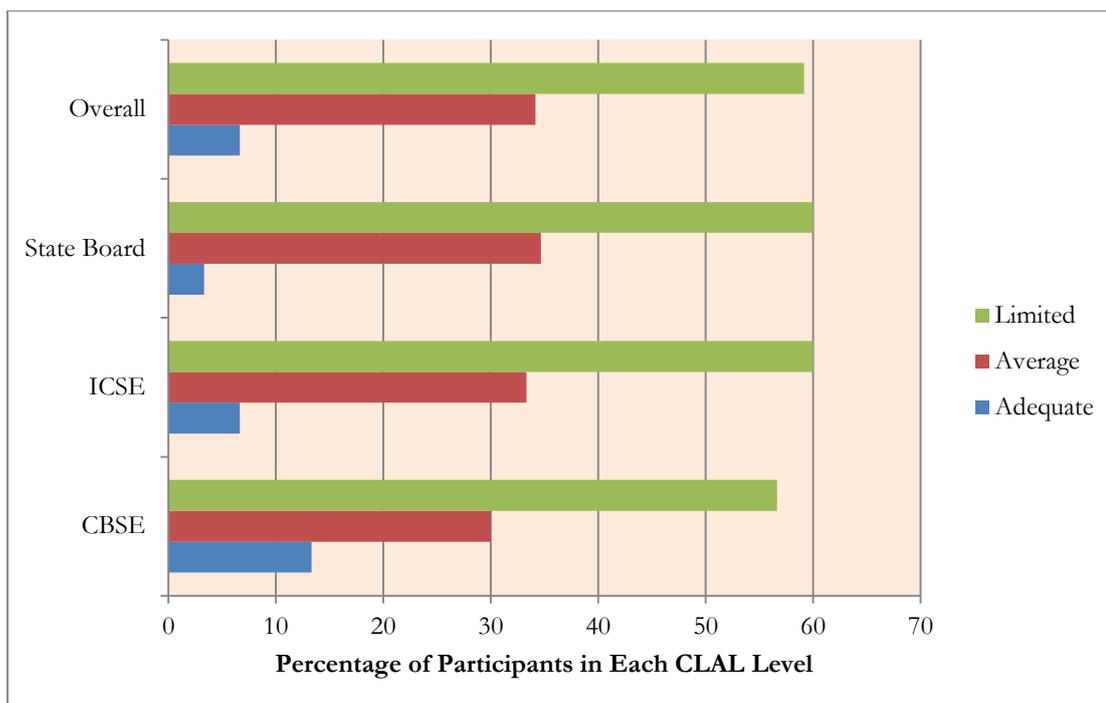
Figure 5.1: Teachers' CLAL levels

Out of the total 120 teachers, 71 teachers, i. e., 59.16% of the total number of participants, were found to have 'Limited' CLAL level. The percentage of teachers falling under the categories- 'Average' and 'Adequate'- comprised 34.16% and 6.66% of the total population. The percentage of teachers in the 'Adequate' category was 13.33 for CBSE teachers, 6.66 for ICSE teachers and 3.33 for State Board teachers. However, the difference in percentage was less in case of the other two categories. In case of 'Average', it was 30% for CBSE, 33.33% for ICSE and 34.66% for State Board respondents. The 'Limited' category got the highest percentage of teachers across the boards of education. Under this category, there were 56.66% of the CBSE teachers and 60% each of the ICSE and State Board teachers.

There was almost a pattern evident from the percentage of teachers found in each level of CLAL. For teachers from three different boards, the highest number of teachers was in the 'Limited' category and the least was in 'Adequate'. This suggests that the CLAL level of English teachers working in schools under different educational boards was far less than what was required to function effectively as teachers. This survey-cum-test may not be a fool-proof test of their CLAL, but considering that the KR 21 (reliability indicator) was calculated to be 0.76 for the survey instrument, the results of the survey can be generalized for a population comprising secondary school English teachers in

Andhra Pradesh. Moreover, this kind of generalisation was necessary from the point of view of this study because it was primarily built on the claims related to the CLAL level of English teachers in the state.

5.4 Intervention to Develop Teachers' CLAL

The result of the state-wise CLAL survey indicated that secondary school English teachers working in CBSE, state-board and ICSE schools had average to low levels of CLAL. When the focus was narrowed down to the three cases, i. e., two secondary level English teachers each from a CBSE, a State Board and an ICSE school, the findings about the teachers' levels of CLAL and practices gave a similar picture. They were found wanting in their ability to construct and carry out classroom assessments. Their beliefs about classroom assessment of language ability, ability to design appropriate classroom assessment tasks and provide constructive feedback to their students, and assessment responsibilities were also analysed. The information obtained from these analyses helped to list a set of areas of language assessment which would help to develop the CLAL of the six selected teachers. While designing the programme and the tasks for the same purpose, the review of existing research literature in the area presented in the second and third chapters of the dissertation was also used for getting guidance and direction.

More details about the programme are presented in the following sub-sections.

5.4.1 Scheduling of the Programme and Rationale

The scheduling of the intervention programme was guided mainly by the principles of Multiple Case Study design. According to those, it is necessary to replicate one study in other sites. Thus, whatever was done at the CBSE school was duplicated at the State Board and ICSE schools (see Chapter 4 for details).

The sessions were integrated into the weekly teaching schedule of the teachers. For each pair, two slots— each lasting around 90 minutes and located during the working hours on week days— were fixed. It turned out to be an effective plan because the researcher trained the teachers on different days of the week. So the impact of fatigue and other such external factors was less on the teachers. Also, as mentioned in the Research Methodology chapter, the intention was to replicate the intervention programme twice. Therefore, it helped that meetings with first pair (CBSE) of teachers were followed by those with the other two pairs.

The researcher met the teachers on the scheduled days with the instruction materials and conducted the TD programme. Though the programme was scheduled to be completed in six weeks, it was completed in eight weeks because some of the pre-planned days were declared holidays by the schools.

5.4.2 Syllabus for Intervention

The broad syllabus for the intervention was finalized fifteen days before the start of the TD programme. The components of the syllabus were carefully chosen so that at the end of the programme, teachers could reach a higher level of CLAL. Though the teachers from each board were expected to carry out assessments of different kinds, their CLAL requirements were more or less the same. The main components included in the syllabus reflected those requirements:

- *language ability*
- *principles of assessment*
- *developing assessment criteria*
- *assessment of individual language skills and components*
- *providing feedback*

Going by the working definition of CLAL proposed at the end of the second chapter, it was assumed that developing the teachers' CLAL, to a great extent, means a sound knowledge about the above-mentioned components and their application in simulated assessment situations. The syllabus can be found as APPENDIX J in the dissertation.

5.4.3 Materials Used for the Intervention

Almost all the tasks used for the intervention programme were interactive, informal and classroom-oriented. While it was made sure to make them challenging for the teachers, effort was also made to ensure that they developed a fair idea about classroom assessment of language in the process. The tasks required the teachers to work individually as well as in pairs and participate in group discussions involving the researcher, who acted as the instructor. The teachers were encouraged to ask questions about the content of learning, especially its utility and application in their assessment contexts, give suggestions, if they had any, to the researcher and make use of their new learning about assessment in their respective classrooms. The researcher monitored and

simultaneously participated in the sessions, and his participation promoted involvement of the teachers in the process. (See APPENDIX K for the tasks.)

Apart from the tasks, the researcher provided the teachers with some reading materials, which included book chapters on language assessment and websites offering useful information about classroom assessment of language ability. The teachers' knowledge about language assessment was kept in mind while choosing those resources. Some of the books, which were utilized, were:

- *Classroom-Based Evaluation in Second Language Education* (1996) by Fred Genesee and John A. Upshur
- *Language Assessment: Principles and Classroom Practices* (2010) by H. Douglas Brown and Priyanvada Abeywickrama
- *The Cambridge Guide to Second Language Assessment* (2012) by Christine Coombe, Barry O'Sullivan and Stephen Stoynoff

The following websites were also part of the reading list:

- http://www.ccsso.org
- http://www.cal.org
- http://www.slideshare.net

5.4.4 Role of the Researcher in the Intervention

As mentioned in the section above, the researcher was actively involved in most of the sessions and worked on the tasks along with the teachers. However, he did not adhere to any pre-determined plan about which role to play at which stage of each session. His intention was to facilitate learning among the teachers. On some occasions, he had to take instant decisions about his role during a session. For example, during one session, the teachers were required to work individually and prepare assessment tasks for various language skills and the researcher had to help each teacher with task preparation. But a couple of teachers did not want to work individually. So the researcher had to let them work as a pair and he just monitored the process without interfering much in it. On some other occasions, when the teachers wanted to discuss, for example, whether to integrate different skills in one assessment, the researcher had to participate in the discussion and think like an English teacher in a school. This kind of flexibility was needed to carry out the intervention effectively and smoothly. It was important to respect and accommodate

the personal preferences, beliefs, knowledge and practices of the teachers while conducting the sessions.

5.4.5 The Sessions

The intervention to develop the teachers' CLAL was organized separately for each selected pair of teachers. The researcher visited each pair on the scheduled dates and at a pre-decided time. The sessions were organized in the teachers' respective schools. The intervention continued for 18 hours spread over eight weeks. The researcher maintained an e-diary and kept an audio-recorded reflective account of the TD programme. Each teacher also kept a reflective journal and noted down their experiences at the end of every week. The following sub-sections present what was done in each week of the intervention and also, an analysis of how the teachers in each case responded to it.

Week 1

In the first week, the focus was on language proficiency, difference between skills and content assessment and importance of teacher-prepared assessment as opposed to external assessments. In all three cases, as teachers had very little training in and idea about language assessment, all the tasks involved some amount of discussion on teachers' existing practices and built on those experiences. Fictional teaching-related accounts were provided to the teachers to help them understand the newly introduced information. After they completed the first five tasks which focused directly on the pre-determined syllabus of the intervention, a reflective discussion was initiated as the sixth task. The teachers had to reflect on their learning and experience in the first five tasks and share their views with the researcher and the other teacher.

Week 2

The basic approach to the intervention did not change much during the second week. Writing assessment objectives and developing assessment criteria— two important aspects of CA— comprised the syllabus for the week. At first, the teachers' understanding of the objectives of their lessons was found out through a discussion and later, there was a discussion on assessment objectives. The teachers were not very sure about the objectives of their lessons. Only after they were introduced to the skills and sub-skills of the language, they seemed to have a better understanding of assessment objectives. Furthermore, it became easy to introduce of the concept of assessment

criteria and enable them to develop their own assessment criteria after that. The teachers could correlate between skills and sub-skills of language and assessment objectives and criteria.

Week 3

The intervention concentrated on availability of resources for assessment, different assessment methods and basic principles of assessment during the third week. It was necessary to make teachers aware of the availability of resources before introducing them to considering different assessment methods. While the task focusing on availability of resources was based on a fictional situation in which a teacher had some difficulty in carrying out an oral assessment in the absence of a voice recorder, the second task focusing on assessment tasks was a mixture of both theory and practice. The teachers were required to use the knowledge and skills they were exposed to in the first and second week when responding to the task. Since the basic principles of assessment have to do with what teachers should do to create effective assessments, they were introduced to teachers in the form of a set of YES-NO questions. In this way, teachers were introduced to new ideas and had the scope to locate their respective practices in the light of the guiding principles. To help them internalize the application of these principles, a list of fictional assessment situations (which they can easily associate with) were given to them in the task and they were asked to match the situations with the corresponding principle.

Week 4

The entire fourth week was devoted to providing the teachers with practice in creating and anlysing tasks for assessing reading, writing, speaking and listening skills along with grammar and vocabulary. They were encouraged to make use of the assessment tasks they had already used for their internal assessment purposes and make improvements in them, wherever necessary. They were provided with a range of options, in terms of content, types of questions, assessment methods, etc., to choose from while developing the tasks. The aim was to develop awareness among them and encourage them to utilize those options in future assessment situations. In addition, the intention was also to inculcate minimum theoretical awareness about assessment task design in them.

Week 5

Two of the important and yet, sometimes neglected, aspects of CA— interpretation of assessment results and feedback— were the focused areas during the fifth week. The tasks, apart from providing teachers with ideas and opportunities to work on assessment scores, aimed to make teachers reflect on their own practices. It was easy for teachers to connect and compare what they practiced with what they did in each task. For enabling them to employ various feedback techniques, they were made to reflect upon given clues, which included questions related to when and how to provide feedback, how much of it, whether oral or written, direct or indirect, positive or negative, individual or group, teacher or peer, etc. This was followed by introduction to different aspects of skills on which feedback can be provided.

Week 6

In the last week of the intervention, the teachers were familiarized with various alternative methods of assessment and how to evaluate and improve already used assessment tasks. Since the teachers had to deal with FAs, they were provided with a few alternative assessment methods through tasks. After introducing them to the meaning, principles and utilities of a few alternative methods, they were asked to select methods that would be suitable for their situation and provide a rationale for their selection. The objective was to enable and encourage them to find solutions to their assessment related problems in alternative methods. Evaluating the effectiveness of the used assessment tasks and making them better formed the second part of the syllabus for the week. Before the teachers were made aware of how to look for areas of improvement and how to improve tasks, a task-based discussion was initiated to trace their practices related and beliefs about evaluation and improvement of used tasks. It was assumed that such a move would lead to reflection on the part of the teachers.

5.4.6 Teachers' Response to the Intervention

Case 1 (CBSE)

Week 1

The researcher had become familiar with the CBSE teachers before he started the intervention. Since one of these two teachers had a better CLAL level, it was easier for her than the other teacher to understand and work on the intervention tasks. Even

during pair work sessions, she took the lead and helped the other teacher understand the requirements of the task. But she was not dominating.

It was important for both the teachers to understand the concepts like "language proficiency", "assessment of language skills and content" and "teacher-based/classroom assessment". These concepts are fundamental to CLAL. Though the teachers found it easy to establish a connection between what they do in their own classroom and the assessment contexts described in the tasks, they had some problem in understanding why they should assess language skills and not content. It was evident from the way *Teacher 2* wrote in his diary:

> *I knew nothing about skill and content assessment. I had a confusion. If I teach my students about Daffodills why I cannot assess their knowledge about it. But I learnt that I must teach and test comprehension skill. Daffodills are mediums only.*

This was an important phase considering how they were made to question their own assessment practices. In the beginning, they resisted. Later, however, they started believing the researcher and accepting the importance of assessing language ability in the classroom. Their response to Task 5 stood out among others. The task contained a set of assessment situations describing the problems faced by teachers when they receive question papers designed by a District Centre or people not working with their students and in their school situation. Both the teachers were quick to relate the given situation to that of their own and in the process, started realizing the meaning and importance of classroom assessment. The first teacher mentioned it in her diary:

> *The word 'classroom assessment' was new to me. The fifth task showed one common problem teachers face. The question paper is a big problem. If the teacher knows how to prepare question paper this problem can be solved.*

One very important point raised by the teacher was teachers' ability to design good classroom assessment tasks. It was quite encouraging because it came from a teacher. The researcher had a discussion with the teachers about what kind of professional support teachers need to carry out classroom assessments effectively.

Week 2

During the second week, the teachers performed the intervention tasks at a faster pace. They took keen interest in the components of the programme. They thought the

programme would help them carry out assessments really well. One of the teachers (*Teacher 2*) wrote about it in his diary:

> *This week was fantastic. I got more support from the instructor. I learnt to set teaching goals. It was a realisation. I want to assess skills only those I have taught. I will divide these skills into parts and give marks to students for each part.*

The teacher's diary entry suggests that he found the tasks useful. Along with his colleague, he worked on connecting assessment objectives with those of teaching. In fact, it is one of the foremost concerns of CA to bring teaching and assessment together. Moreover, this realization is an important part of CLAL.

The teachers also worked on tasks aiming to promote understanding of assessment criteria. The researcher participated in the sessions and worked with both the teachers to develop a set of assessment criteria. It had a positive impact on teachers. One teacher (*Teacher 1*) pointed out how she felt about it:

> *I was happy to see that this week's tasks were directly connected to my teaching. The trainer also solved those tasks with us. It was a bonus. We discussed everything in the group...It was a great achievement when we developed the assessment criteria.*

It was also noticed that the teachers got deeply involved in the tasks related to the development of assessment criteria. They requested the researcher to get them each a copy of the handout containing the list of sub-skills. They wanted to make use of the same list in their future classroom assessments. As they were already practising CCE in their curriculum, they found these tasks directly relevant to their needs.

Week 3

The focus was on assessment methods and basic principles of assessment during the third week. These two components are the pillars on which the field of assessment stands and high level of CLAL also means a thorough understanding of and ability to apply these concepts. The teachers did not have any problem in relating the tasks of the previous weeks to those of the third week. However, they struggled to apply the concepts to the real life assessment situations. They took a lot of time to complete the tasks, especially the second teacher (*Teacher 2*). Since there was constraint of time, it was not possible to spend more time on each component. The other teacher (*Teacher 1*) rose to the occasion and took the responsibility of discussing the concepts with her colleague

whenever they had time outside the sessions. She explained to him the meaning of the continuum of assessment methods ranging from indirect to direct though the researcher did not initiate this collaboration. However, the researcher provided multiple examples of different kinds of assessment methods. But the teacher continued to struggle in the third and fourth tasks. The other teacher also asked the researcher to help her with matching the assessment situations with the corresponding principles of assessment. When the researcher made use of a few leading questions, things became substantially easier for the teachers. Both the teachers even mentioned this in their diaries:

> *The tasks for this week were more difficult than last week. The concepts were new. But I need to know to become a good teacher. I can use CCE well then only. ...I could not understand direct and indirect assessment at first. But the trainer asked us very small questions and guided us. I went home and read the tasks again. Then only I understood more clearly. (Teacher 1)*

> *I worked much more this week. I looked at all the tasks at home and tried to understand them better. The first task was the only easy task. The last task was very hard. I could not match five situations correctly. But the trainer asked me some simple questions. I got my own answers. I liked that trick. (Teacher 2)*

Week 4

The main focus of the fourth week was on development of tasks for assessing language skills and components. The teachers made note of almost everything while taking part in the programme. They thought they could immediately put the newly acquired knowledge and skills to practical use. They were required to work and generate tasks with the help of given clues. Both the teachers managed to design fairly acceptable tasks.

One (*Teacher 1*) of the teachers had recently undergone some training in CA. She took the lead and tried to mix her own ideas with those of the other teacher for developing and fine-tuning the assessment tasks. But she informed the researcher that she was not conscious about several factors like text types, difficulty levels and wide range of purposes of using language skills. She included it in her reflective diary:

> *I attended a training programme recently. It was on CCE. The training was on assessing oral skills. I learnt a lot there. But I got more to learn last week. The*

> *information about text types and difficulty level of tasks enriched me. Every English*
> *teacher should know about this.*

The other teacher liked the tasks too. He thought the steps which he was asked to follow to design tasks gave him the direction he wanted. He had the opinion that such information should be shared with other teachers so that they can put the information to use for task design. It would make CA more effective.

Week 5

The syllabus for the fifth week of intervention comprised analysis and interpretation of assessment results and providing constructive feedback to students. Both the teachers enjoyed the arithmetic calculations mentioned in the hand-outs. However, they were slow in interpreting scores. So they took extra time to complete the tasks and even for that, the researcher had to guide them at every stage and make them think using some hints and leading questions. In her diary, *Teacher 1* mentioned this thing:

> *I never imagined that students' marks could have so much meaning. In the beginning I*
> *thought it will be easy but it was so difficult. Both I and my colleague had tough time.*
> *But the instructor was kind enough to help us. His clues helped us to find answers.*

The teachers also found it challenging to perform the task involving two skewed graphs (Task 2, 1. C) representing students' scores. But it was the tasks on feedback that was found more interesting by *Teacher 2*. He was surprised to see that teachers should think about several things while offering feedback. He liked the first task on feedback a lot:

> *I enjoyed the feedback tasks. It was a complete new thing for me. I did not know about*
> *so many factors in feedback. The first task about feedback taught me many things. I*
> *feel I can now give good feedback to my students.*

Week 6

During the last week of intervention, the teachers participated actively in the sessions. They seemed to enjoy performing the tasks. In addition, they got introduced to application of alternative methods of assessment and evaluation of assessment tasks. They had heard about portfolio but did not know much about it. They knew little about other alternative methods of assessment. The first teacher talked about it in her diary:

> *I knew about portfolio from a teachers' manual but my understanding was limited. I did not know how to make use of portfolio, journals, interviews and others for class tests. The tasks not only introduced these new things but also gave me more options for formative assessments.*

Also, the teachers were made aware of evaluation of assessment tasks. The set of criteria which was used in *Task 3* made things simple. It helped teachers to reflect upon their own assessment tasks while evaluating that of others. The second teacher wrote in his diary that he wanted more such tasks for practice:

> *The task about task evaluation was an important thing. Every teacher should learn to do it. But we need more practice. I understood the concept but want more practice.*

Case 2 (State Board)

Week 1

Like the first case, the researcher met the two State Board school teachers much before the intervention. The teachers were quite friendly and had a positive attitude towards the programme right from the start though their CLAL levels were on the lower side. So it was necessary to support them a little more during the initial stages of the programme. They faced some amount of difficulty in understanding the task contents. The researcher had to monitor their responses and ask them a few leading questions to guide them and help them understand concepts like "objectives of teaching", "successful language learner" and "relationship between examination performance and ability to use language effectively" covered in the first and second task. A similar strategy was also adopted for the next task in which differences between assessment of skills and content, and classroom assessment and standardized testing were discussed. The teachers found post-task discussions, in particular, quite useful. One of the teachers (*Teacher 3*) talked about it in her diary:

> *I had some doubts when I was doing the tasks. But I got answers to my questions during the discussions. We asked questions. I wanted to gain more knowledge. So I asked many questions. I can say these discussions were useful and I request the instructor to discuss more.*

But the fourth and the fifth tasks posed more challenges to this pair of teachers. The researcher planned to ask the teachers to work individually in the first phase of the fifth

task but later, he had to change the plan and let the teachers work as a pair. They helped each other and explained to each other the task requirements in Telugu, their mother tongue. The researcher played the role of a facilitator and just ensured that the purposes of the tasks were achieved to the maximum possible extent. The sixth task was used for consolidation. One teacher (*Teacher 4*) seemed to like it:

> *This week I learned so many new things. It was not bed of roses. After a gap I am fortunate to get this opportunity. After everything, I got clarification from the instructor. The last task was full of discussion. It was enlightening. I got confidence to express myself.*

Week 2

The teachers also showed a lot more enthusiasm during the second week of the programme. They stayed back after the scheduled time for the sessions and discussed assessment-related problems with the researcher. As the state-board follows an approach similar to CCE, the teachers found the tasks relevant to their needs. They took interest in knowing the sub-skills and developing assessment criteria. Though they struggled a little to define levels of achievement of learners in the beginning, they found it much easier when the researcher showed them by defining levels of language ability and the components they comprise. The researcher tried to make them see how they could use the list of sub-skills for developing the criteria. One of the teachers (*Teacher 3*) wrote about it in her diary:

> *I just heard about assessment criteria before. I have now some knowledge about it. It is new thing and very useful for me. We can utilise sub-skills when we prepare assessment criteria in our school... Through tasks only I got this idea.*

The teachers had to work hard to complete the fourth and fifth tasks of the week. Although they worked as a pair, they could not write the descriptors well. But they kept trying. The researcher, while discussing their task performance, helped them by raising a few leading questions like:

- What can be an appropriate description for "excellent organization" in writing?
- How will you define the next level of ability in the same category?

At one point, the teachers wanted the researcher to write down the descriptors. It was quite understandable because they were perhaps not very comfortable with their extended struggle.

Week 3

The tasks of this week posed two different kinds of challenges for the teachers. On the one hand, they had problems with the new concepts and on the other, they had to read and respond to the tasks based on these tasks. This dual challenge was a little too daunting for a teacher (*Teacher 4*). He felt demotivated and could not hide his frustration in his diary:

> *I think the tasks were very difficult this week. I could not understand most of them.*
> *The methods task was easy. The fourth task was complex. I wanted more examples*
> *like that task. I never practised that kind of matching task. So I did not do it*
> *properly. The trainer should make it easier why because we are just school teachers.*

In contrast, the other teacher (*Teacher 3*) accepted the challenge with a positive mindset. She did not mention any kind of difficulty in her diary. She thought challenging tasks could help her learn about different aspects of assessment thoroughly:

> *The second, third and fourth tasks were tough. But I think I can acquire more*
> *knowledge if the tasks are tough. The trainer's presence made the difference. He guided*
> *us. I never lost my way and deviated. I clarified all my doubts then and there. I would*
> *not use my brains if I had easy tasks.*

The differences in thinking between the two teachers suggest that every teacher is an individual and individual differences should be respected while educating teachers. It also implies that the teacher educator needs to adjust his/her roles depending on the individual or group of teachers he/she is handling.

Week 4

The teachers could connect the intervention sessions and the tasks used during the week to their assessment responsibilities. They were happy about knowing the steps using which they could construct classroom assessment tasks. The body language of the

teachers was quite positive. They had a lot of excitement while working on tasks. One of the teachers (*Teacher 3*) was all praise for the hand-outs:

> *The tasks brought enthusiasm in me. But the hand-outs were the real subject of interest. I have saved all the hand-outs in a folder. I will use them whenever we have formative tests. They are like a white stick for a blind man like me. I must thank the trainer for sharing this valuable information with us.*

It was very obvious that they wanted to know about task design and the tasks provided much of what they wanted. The tasks turned out to be almost like self-learning materials. In this connection, the second State Board teacher (*Teacher 4*) shared his experience in his weekly diary:

> *I want to know more about CCE and how to make good tests. So I liked this week's tasks. They were more practical. I had discussed about using this information with my colleague. They were easy and interesting like distance education materials. So I did all the tasks myself.*

Week 5

The teachers faced a great deal of difficulty in responding to the tasks of the week. They did not enjoy calculations. Both the teachers asked many questions to the researcher about why they should learn to calculate and interpret the scores. They were convinced that if someone works hard, they can score well. So why 'play' with scores. The third teacher did not hide her feelings in her diary:

> *I was confused about interpreting scores. I believed that students can do well if they work hard. Then why I should play with their scores. But now I feel the scores can help to give feedback to students and evaluate our teaching and examination question paper.*

But she found the feedback part immediately useful. She thought the sessions could improve her ability to provide constructive feedback to her students. Even *Teacher 4* had a similar belief. He had difficulty in accepting that he needed to interpret his students' assessment scores. But he was impressed with the tasks on feedback. He kept the hand-outs for future references.

Week 6

During the sixth week, the teachers had no problems in understanding the task requirements and content. However, they had the apprehension that the alternative methods were not really practicable because the school administration would not allow such practices. One of the teachers (*Teacher 4*) pointed this out in his diary:

> *It was very interesting to know about portfolio, interview, observation and questionnaire. But I don't think other teachers and headmaster will allow me to use these things for formative tests. There are also some practical problems like students will not be serious about tests.*

The third teacher did not think differently about alternative assessment methods. But both the teachers found the task evaluation part something necessary to learn. They paid a lot of attention to the criteria of evaluation. In her diary, *Teacher 3* shared her thoughts:

> *I am 100% sure that task evaluation is necessary for teachers. How will they improve if they don't evaluate? Teachers should work in a group and evaluate each other's tasks. Then only it will be fruitful. For that purpose, they need to get this kind of information.*

Case 3 (ICSE)

Week 1

Though linguistically the two teachers of the third school (ICSE) were better than the first two pairs, there was not much gap between their CLAL levels. They had deep-rooted beliefs about assessment which were not easy to mould or change. They openly asserted that they did not need the TD programme. It was a huge challenge for the researcher to prepare them for the sessions. The diary entry of *Teacher 5* provided more information:

> *For me it was rather quite strange that we could test students' progress without examinations. There are, of course, individual views about this. But personally, I feel students must pass a tough test and prove that they have learnt the lessons well. Teachers can be biased and give marks to their students. So examinations are necessary whether one likes it or not.*

Though they accepted that classroom assessments can contribute to students' learning, the researcher could not convince them that such assessments can and should replace centralized examinations. They also thought that by diluting the importance of examinations, the government was spoiling the future of children. It was unacceptable to them that learning should and can happen without fierce competition. The second teacher (*Teacher 6*) wrote about it in her diary:

> *This training programme did not teach me anything new. The promotion of CCE at the cost of examinations is a farce. I'm not sure if it will help the education system. We need to ask this question— do students know what is good and what is bad for them. Even we adults don't. Let the experts take care of question papers. Aping experts won't open new vistas. I thought what we did last week was exactly that.*

At every stage of the intervention during the first week, the two teachers tried to resist aggressively. But the researcher had already anticipated such reactions from them. It was not unusual. At the end of the week, they were made to face many questions about their own assessment practices.

Week 2

The teachers responded better to the intervention tasks than the first week. Though they had already informed the researcher that they were aware of sub-skills of major language skills and assessment criteria and that they did not need any professional support in those, still they found it difficult to write the level descriptors and define different components of assessment criteria. They softened their attitude afterwards and accepted help from the researcher to complete the tasks. One teacher (*Teacher 5*) indirectly talked about her experience in the diary:

> *The second, third, fourth and fifth tasks were not new for me. I have been using similar approaches in my class. But the terms were new. The variety of descriptors and descriptions needed a fair bit of hill-climbing. ...I had no qualms accepting help from the trainer's side.*

The same teacher suggested that ready-made assessment criteria should be made available to teachers in addition to training in developing such criteria. In fact, she claimed that not all teachers have time and opportunity to undergo training. So it would be better if they were directly supported with flexibly-designed assessment criteria and scoring rubrics.

The teachers did not face any linguistic difficulty, but it was obvious that they found it challenging to relate the concepts with their own assessment situations. They were given a few prompts in the form of questions during the third task and they used the same to perform the fourth task. They had less problems than the other groups in matching the situations with the principles of assessment presented in the fourth task. The researcher did not have to intervene much. The teachers discussed the situations and the principles of assessment with each other. One teacher (*Teacher 5*) wrote about it in her diary:

> *I'd say this week was more satisfying than the previous weeks in terms of learning about real assessment. My colleague was more than eager to discuss the tasks. The tasks were a little tricky but not impossible to complete. The question-prompts used by the trainer for the third task was interesting. I am planning to use them in my class.*

The other teacher (*Teacher 6*) had a very similar opinion as she continued to follow her mentor and guide. She talked about it openly:

> *I thoroughly enjoyed working on this week's tasks with my senior colleague. The amalgamation of useful tasks and a person ripe with experience sitting next to me made it a great experience. I hardly faced any hurdles. To be honest I enjoyed the last task. I had fun.*

It was a positive change on the part of both the teachers though they had problems initially in accepting that they needed some orientation in CA. However, they were less rigid and more cooperative than the previous weeks. The researcher got some encouragement from this slight change in their behaviour. It helped in the smooth transaction of the programme.

Week 4

Though these two teachers did not seem to be as enthusiastic as the teachers from the other schools, they certainly gave the impression that they liked the tasks. They asked quite a few questions of the researcher about the information provided on the hand-outs. Most of these questions were related to the importance of knowing the steps in task-construction and the scope of the frameworks provided to them. Both the teachers seemed to be convinced with the researcher's explanation. It was found in the diary entry of one (*Teacher 5*) of the teachers:

After getting convincing explanations about the frameworks of the task design, I'm now confident about testing my own students. The hand-outs provided by the trainer will help me. They have given us a perspective about assessment and will work for us as touchstones when we design assessments in future.

The other teacher too had similar thoughts. She thought they could learn a lot from the answers given by the researchers to their questions. She also mentioned in her diary that every training programme for teachers should have a question-answer session in which teachers would be allowed to pose questions to the trainer and so that in the process, they could acquire more professional knowledge. She also suggested having more such talks, as happened during the week, for the next two weeks. She specifically mentioned tasks used in the listening section.

Week 5

The teachers from the ICSE school showed some interest in the tasks used during the fifth week. The sixth teacher responded to the tasks on interpretation of scores better than the fifth teacher. She made use of the hints offered by the researcher well and answered most of the questions correctly. She also participated in the discussions enthusiastically. On the contrary, the fifth teacher was slow to respond to the tasks. She was unsure about how interpretation of scores would give new insights about students' performance and the utility of tasks. She wrote about it in her weekly diary:

At the outset, I had just one query: How will it help me in improving my students' language skills? By the time I had completed the tasks, I realized that I could figure out quite a few things about my students' progress from their scores.

The teachers did well in the second set of tasks focusing on feedback. They had little problem in understanding the task requirements. They made use of information presented through *Task 3* while performing the fourth task. Moreover, they initiated discussions while responding to *Task 4* and involved the researcher. The sixth teacher found it a good experience:

I always believed that discussions are a good way of learning. Our discussions on feedback and the amazing amount of information we had about it gladdened my soul. I was elated to have learnt so much about important things in language teaching like feedback.

Week 6

The teachers had similar response as State Board teachers to the tasks based on alternative assessment methods. Both the teachers told the researcher that those methods might be interesting and useful but the ICSE assessment policy does not offer the required flexibility to use methods like portfolio and observation. Even though the researcher tried to convince them about using these methods for promoting learning, they still had a lot of doubts. It was evident from the diary entry of *Teacher 6*:

> *On the one hand, it is important to learn about these methods, on the other hand, the practical constraints must be thought about thoroughly. I have little disagreement with the instructor that these methods will help students learn better. But the principles of the board and the school don't allow that freedom.*

However, the teachers showed a very positive attitude towards task evaluation. They were very convinced about it and wanted to make use of it at the earliest opportunity. They planned to do it with other teachers. The fifth teacher showed her interest in the activity and thanked the researcher for introducing her to the concept. She confessed that she never thought about it though the concept did not sound unfamiliar to her.

5.5 Changes in Teachers' CLAL Level

As per the research design, the impact of the TD programme on the CLAL levels of the teachers under the three cases was calculated by tracing the changes between the teachers' pre- and post-intervention tasks. A level was assigned to each teacher according to their total score on the instrument (see section 5.3 and Table 5.2 of this chapter). To maintain clarity and bring effectiveness to the analysis of the teachers' performance on individual areas and aspects of assessment, the items in the survey instrument were divided into seven broad sections and a few basic components of assessment were included under these sections:

Table 5.4: Sections in CLAL survey instrument

Broad Section	Question Number in the Instrument	Components of Assessment
1. Reading	1	validity, reliability, practicality, authenticity, interpretation of scores, task evaluation
2. Writing	2	reliability in scoring
3. Speaking	3	validity, authenticity, formative assessment, observation
4. Grammar	4	Authenticity
5. Feedback	5	feedback, rater reliability
6. Classroom Assessment	6	learner- and learning-friendly assessment, alternative assessment, wash back, diagnostic assessment, authenticity
7. Plan for Classroom Assessment	7	

The framework above was kept in mind while analyzing and calculating the responses of individual teachers to questions in each section. The teachers' pre- and post-intervention CLAL levels and the corresponding changes are discussed in the subsequent sections under three cases. In addition, an overview of changes for each case is also presented. Finally, a cross-case analysis is carried out.

5.5.1 Case 1: Pre-, Post-Intervention CLAL Levels and Changes

Pre-intervention CLAL Levels

Teacher 1

The first teacher scored 22 out of 38 and thus, she was assigned a 'Limited' CLAL level. Her scores in the seven major sections are presented below:
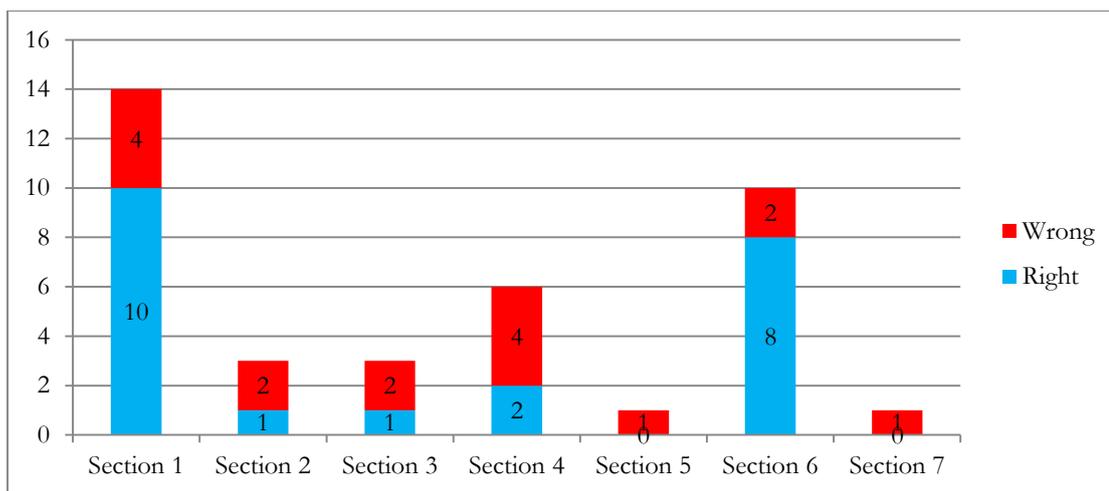
Figure 5.2: Performance of _Teacher 1_ on CLAL instrument

As shown in the above figure, she performed better in sections 1 and 6 than in other sections. In other words, she showed some understanding of the basic principles of assessment in relation to Reading and even better grasp of Classroom Assessment. But she seemed to be unsure about task objectives and involving students in the process of assessment.

In the other sections, her performance was less than average. She got it completely wrong in sections 5 and 7 which focused on feedback and plan for classroom assessment respectively. However, in sections 3, 4 and 2, she managed to get 33.33% of her answers right. This indicates that she had problems in responding to questions related to Speaking, Writing, Grammar and aspects of assessment like rater reliability, formative assessment and authenticity.

Teacher 2

The second teacher got 21 on the CLAL instrument. So he was placed in the category of one with 'Limited' level of CLAL. His performance in different sections of the instrument is presented in the following diagram:
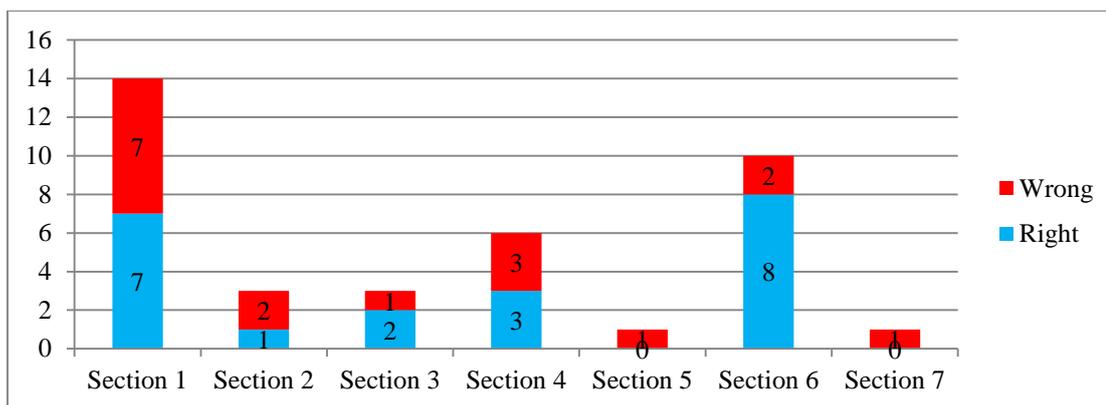
Figure 5.3: Performance of *Teacher 2* on CLAL instrument

The diagram above shows that he performed better than average only in section 6. In the other six sections, he did not fare well. This means that he had a fair idea about classroom assessment and some amount of knowledge about assessing writing, reading and speaking skills and aspects of assessment like authenticity, informal assessment, rater reliability, assessing grammatical ability and authenticity. Like his colleague, he did not respond correctly to questions on task objectives and involving students in the process of assessment.

His performance was mediocre in sections 5 and 7 which focused on feedback and plan for classroom assessment respectively. It could be deduced that he had problems in understanding how to create appropriate rubrics, maintain reliability in scoring, provide feedback and plan a classroom assessment effectively.

Post-intervention CLAL Levels

Teacher 1

The following figure shows the performance of *Teacher 1* on the CLAL instrument.
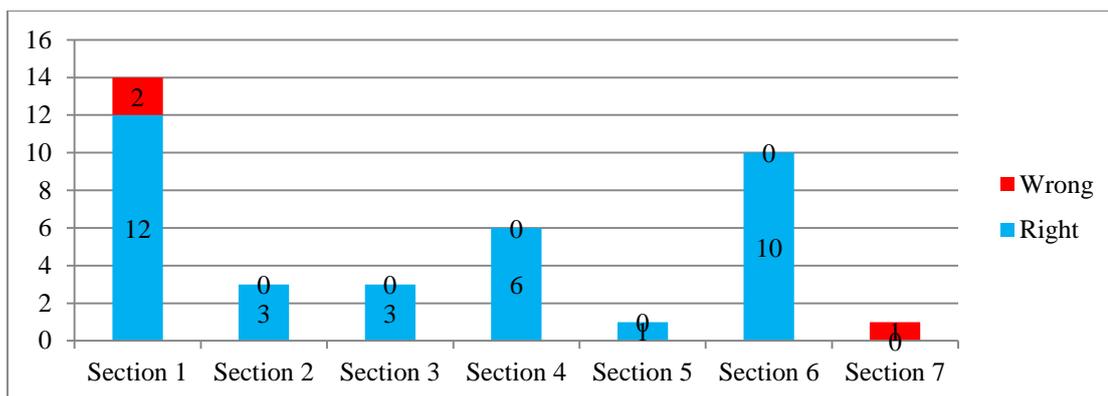
Figure 5.4: Performance of _Teacher 1_ on CLAL instrument

The teacher got 35 out of 38. This high score indicates that the teacher acquired the 'Adequate' CLAL level. Her performance in all the sections except the last one was very good. She displayed excellent understanding of concepts related to classroom assessment, authentic assessment of grammatical ability, assessment of writing, speaking and reading and rater reliability.

Teacher 2

The second teacher scored 34 on the instrument. His performance in the the individual sections of the instrument is represented in the chart below.
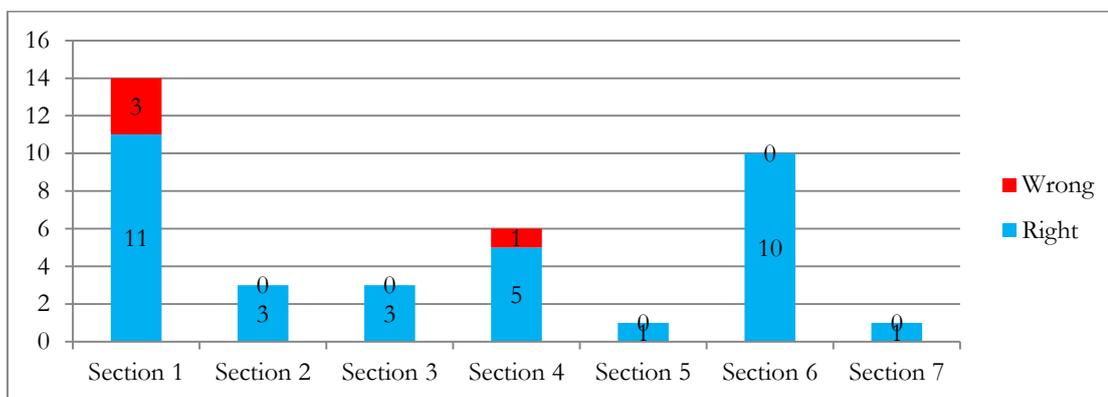
Figure 5.5: Performance of _Teacher 2_ on CLAL instrument

In sections 2, 3, 5, 6 and 7 the teacher got all his responses right. However, in section 1 he got eleven responses right and in section 4, he got five out of six responses right. Considering that the percentage of the wrong responses was around 10%, the teacher's performance can be regarded as very good . Like his colleague he did very well in sections focusing on classroom assessment, writing, speaking, feedback and plan for

classroom assessment. Even in sections 1 and 4 which focused on assessment of reading skills and grammatical ability, he scored around 78% and 83% respectively.

Change in Teachers' CLAL

Teacher 1

This teacher scored 22 and 35 on the CLAL survey instrument before and after the intervention respectively. The percentage of improvement was 34.21. Her performance in the individual sections of the instrument is presented in the following diagram. Apart from the section on plan of assessment, all other sections saw some improvement.

	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Pre-intervention	71.4	33.33	33.33	33.33	0	80	0
Post-intervention	85.71	100	100	100	100	100	0
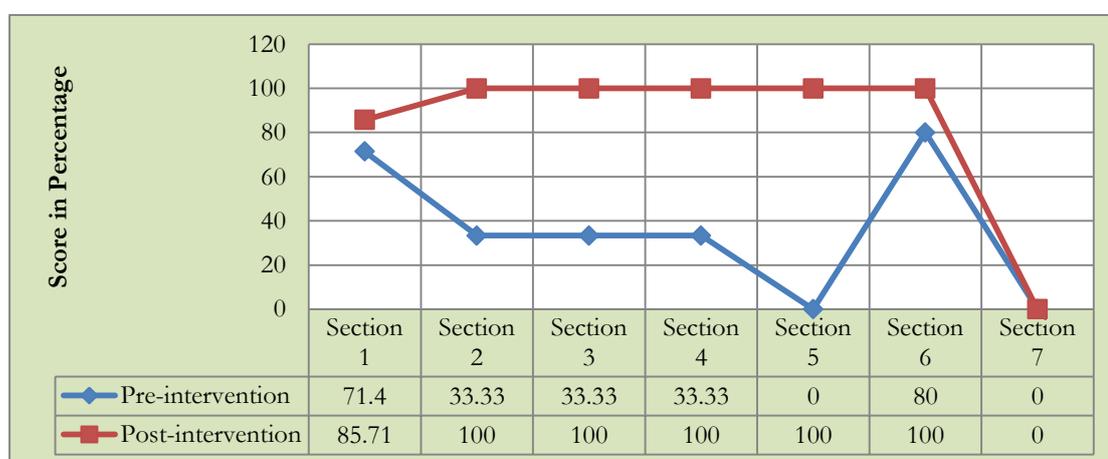
Figure 5.6: Pre- and post-intervention performance of *Teacher 1 on the* CLAL instrument

The percentage of increase was highest in the feedback section followed by the sections on writing, speaking and grammar, which saw an increase of more than 60%. The surge in CLAL score indicated that the teacher understood the components of CA presented during the intervention programme well. At least, she broadened her CA knowledge-base through the TD programme. It was a result of her self-motivation to grow professionally, interest in the programme content, ability to read, discuss, understand and analyse assessment-related issues and the need-based and intensive nature of the programme. Though she could not answer the last question on the instrument correctly on both occasions, in all other sections, she exhibited progress: her overall post-intervention CLAL score showed remarkable improvement.

Teacher 2

The CLAL level of *Teacher 2* got enhanced after undergoing the TD programme in classroom assessment which lasted for eight weeks. The teacher, who scored 34 on the post-intervention, had a score of 21 in the pre-intervention administration of the CLAL survey instrument.

	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Pre-intervention	50	33.33	66.66	50	0	80	0
Post-intervention	78.57	100	100	83.33	100	100	100
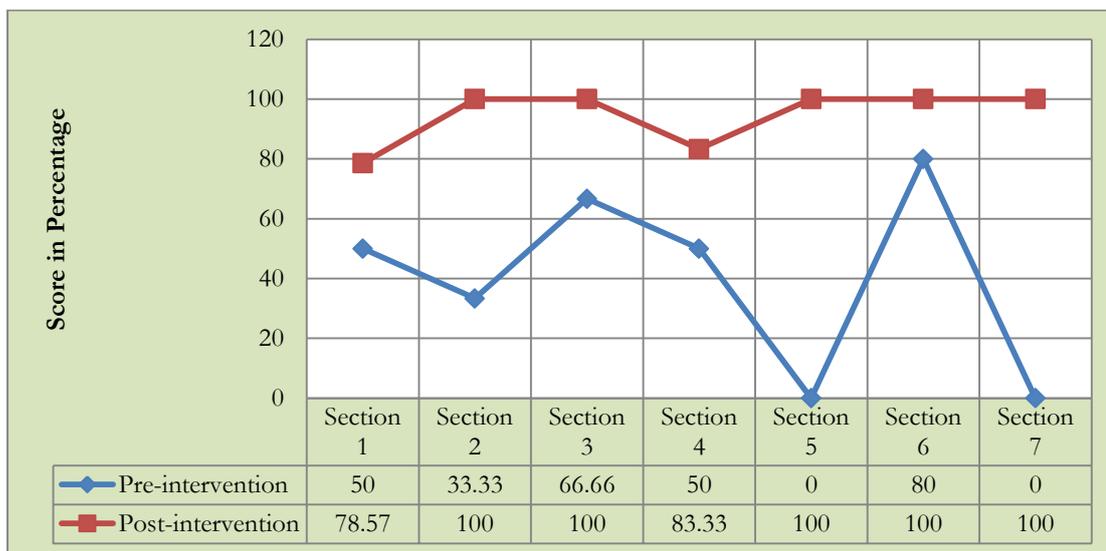
Figure 5.7: Pre- and post-intervention performance of *Teacher 2 on the* CLAL instrument

The teacher displayed an increase of 34.21% in his CLAL score. He bettered his pre-intervention score in all the sections of the instrument. While 100% improvement was recorded in the sections on feedback and plan of assessment, the score for the section on writing went up by more than 60%. The scores for all the other sections also went up with the classroom assessment section marking least percentage of increase. Though he was a little less motivated than his colleague and did not immediately put the newly gained knowledge into practice, nevertheless, he performed well on the CLAL survey instrument after the intervention. The reasons for his good performance could be his engagement with the tasks used during the sessions, his realization about the utility of the TD programme for his professional growth and the programme meeting some of his immediate professional needs.

Overview of Changes in Case 1 (CBSE)

Both the Case 1 teachers showed 34.21% increase in their CLAL scores. In the sections on writing, feedback and CA, their progress was identical. Some difference was found in

sections on reading, speaking and grammar. The second teacher got the only question in the last section right whereas the first teacher got it wrong. Overall, the CBSE teachers, who had similar CLAL level before and after the intervention, made a fair progress after attending the TD programme. However, their understanding of the concepts like that of validity, reliability and authenticity did not seem to improve much. The reasons could be many. The teachers might need more time to understand and apply these concepts. Moreover, the small number of questions for individual sections in the CLAL instrument may not give a clear picture of the improvement.

5.5.2 Case 2: Pre-, Post-Intervention CLAL Levels and Changes

Pre-intervention CLAL Levels

Teacher 3

The third teacher scored 22 on the CLAL instrument.
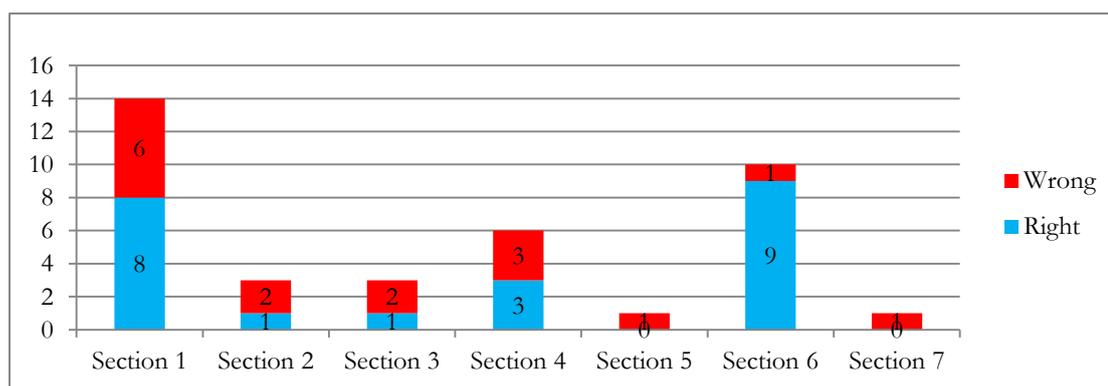
Figure 5.8: Performance of *Teacher 3* on CLAL instrument

The marks scored in each section of the instrument are diagrammatically presented above. The teacher had a 'Limited' level of CLAL. It was also evident from her performance in all the sections of the instrument except the sixth one. She got 90% of her answers right in section 6. The next best performance was in the first section in which she had a score of around 57%. In all other sections, she scored 50% or less than that.

Her scores in individual sections suggest that she had a very good idea about classroom assessment. She also showed some evidence that she was aware of a few things about how some of the basic principles work in the assessment of reading skills. Though she scored 50% in section 4, her answers to individual questions indicate that she was confused about the concept of authenticity.

Teacher 4

With a score of 16 on the CLAL instrument, *Teacher 4* was placed in the category
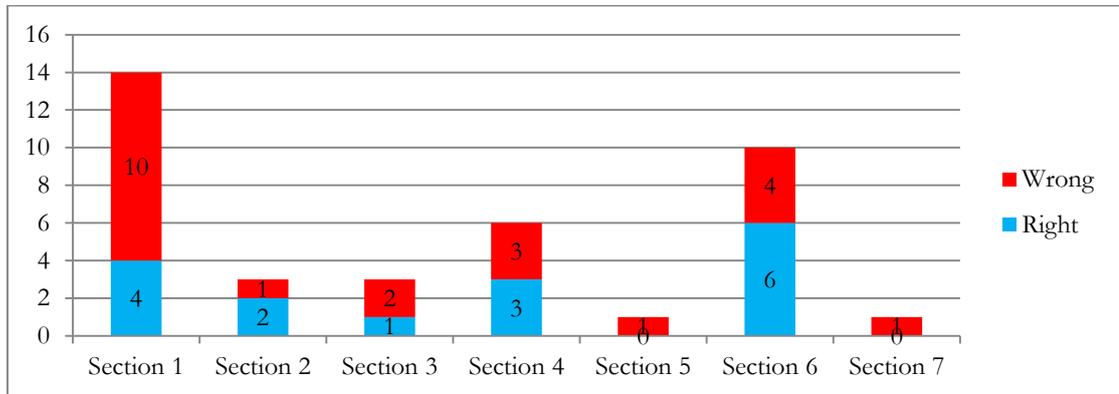
Figure 5.9: Performance of *Teacher 4* on CLAL instrument

of 'Limited'. His performance in sections 2 and 6 was just around average. He scored 66.66% and 60% in these two sections respectively. In sections 1, 3 and 4, he got low scores, and in sections 5 and 7, he got both the answers wrong. As evident from the above figure, his knowledge about classroom assessment and assessing Writing was better than that in assessment of Reading, Speaking, Grammar, giving feedback and planning a classroom assessment.

Post-intervention CLAL Levels

Teacher 3

The following diagram highlights the performance of *Teacher 3* on the CLAL instrument.
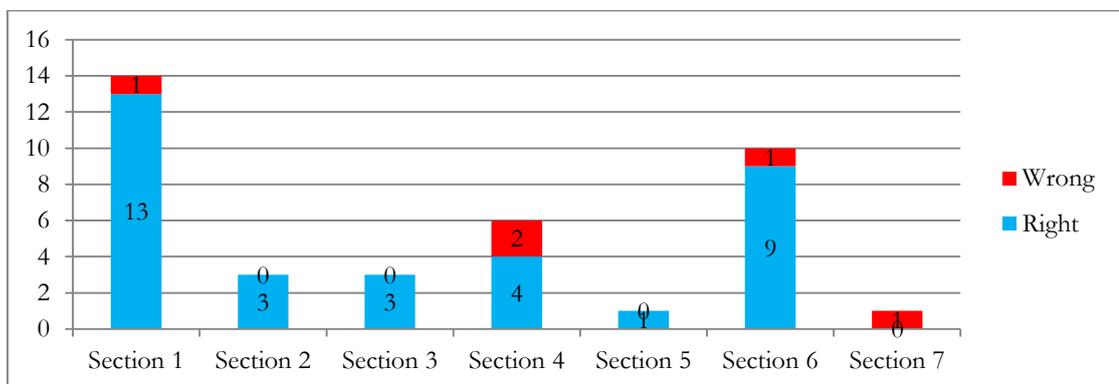
Figure 5.10: Performance of *Teacher 3* on CLAL instrument

The teacher scored 32 and thus can be placed in the category of 'Adequate' CLAL level. She got all her responses right in the sections 2, 3 and 5. Even otherwise, she got excellent scores in sections 1 and 6. Her lowest score was in 66.66% which she got in the fourth section. The scores above indicate that she had deep understanding of basic concepts related to assessment of reading, writing, speaking, providing feedback and classroom assessment. She seemed to have good knowledge about the basic principles of assessment, formative assessment, diagnostic nature of assessment and alternative methods of assessment.

Teacher 4

The following diagram exhibits the performance of *Teacher 4* on the CLAL instrument.
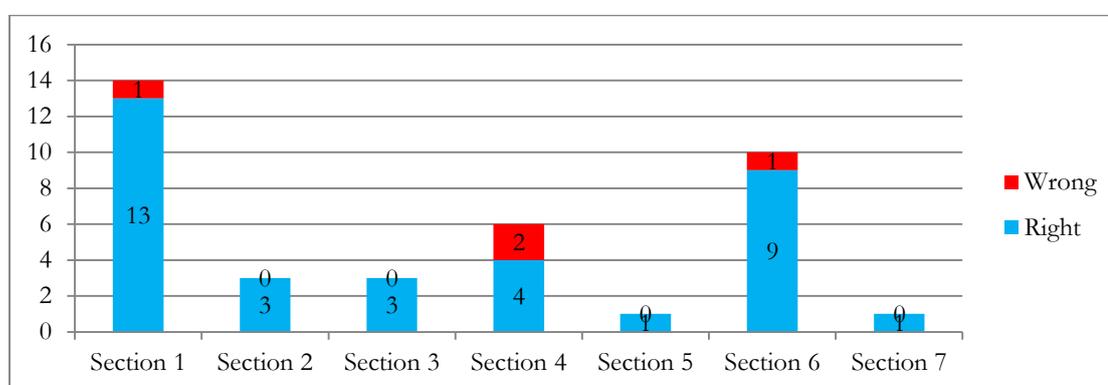
Figure 5.11: Performance of *Teacher 4* on CLAL instrument

The teacher got 34 of her responses right. She got all her answers right in sections 2, 3, 5 and 7. She had one wrong response each in sections 1 and 6 and two, in section 4. These scores reflect his awareness about application of basic principles of assessment to assessing language skills, classroom assessment, alternative assessment and plan for designing classroom assessments.

Change in Teachers' CLAL

Teacher 3

The third teacher, like the first two teachers, also showed progress in terms of CLAL. Her post-intervention CLAL score— 32, was 26.31% higher than 22— her pre-intervention CLAL score. Her performance on the individual sections of the CLAL is presented below diagrammatically.

	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Pre-intervention	57.14	33.33	33.33	50	0	90	0
Post-intervention	92.85	100	100	66.66	100	90	0
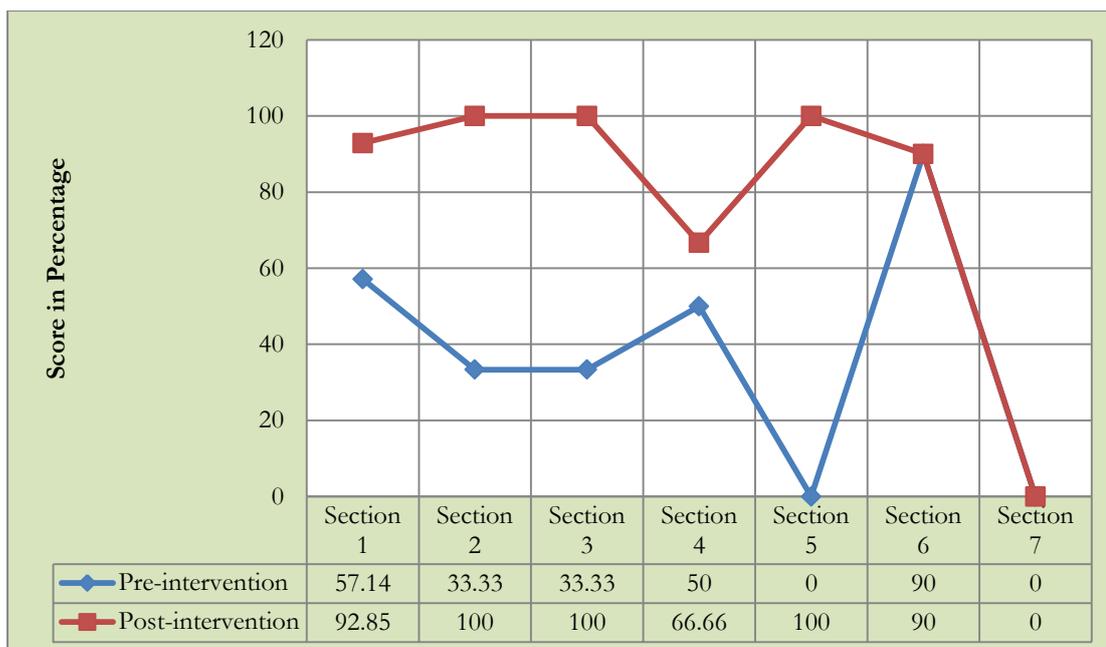
Figure 5.12: Pre- and post-intervention performance of *Teacher 3 on the* CLAL instrument

Apart from section 7, in which she did not score anything on both occasions, in all other sections she showed some improvement. The 100% increase in the score in the feedback section was followed by more than 60% increase in those of writing and speaking. While the reading sections recorded a growth of around 35%, the score for the grammar saw only 10% upsurge and that for the CA section remained the same. The overall gain indicates that the intervention programme, especially the discussions on some of the concepts like feedback, validity, reliability, etc., had productive impact on her CLAL. Since she was unhappy with the fact that the government did not provide adequate training in assessment to teachers, she might have liked the programme for fulfilling her professional requirements related to language assessment. Moreover, she was involved and quite active during the intervention. The above factors could have contributed to the development of her CLAL level.

Teacher 4

The progress made by *Teacher 4* in terms of CLAL was higher than all other teachers. The difference between 16 and 34, i.e., his pre- and post-intervention CLAL scores respectively, was 47.36%. His scores in individual sections are converted to percentage and presented below.

	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Pre-intervention	28.57	66.66	33.33	50	0	60	0
Post-intervention	92.85	100	100	66.66	100	90	0
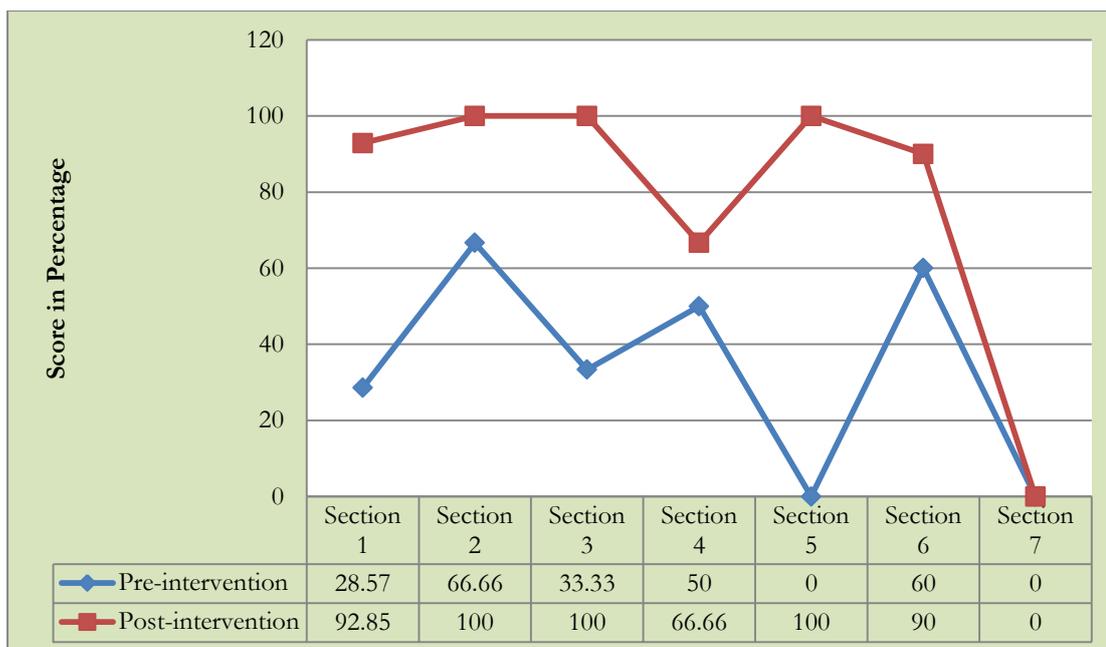
Figure 5.13: Pre- and post-intervention performance of *Teacher 4 on the* CLAL instrument

As evident from the above figure, the fourth teacher exhibited remarkable increase in his scores for most of the sections in the CLAL instrument. Though the teacher could not answer the question on plan of assessment correctly, he recorded 100% improvement in the feedback section. Also, positive changes in the scores were recorded in the sections on reading, writing, speaking, grammar and CA. It can be guessed that the discussions on sub-skills and the basic principles of assessment during the intervention session might have contributed to this upsurge in scores. His strong intrinsic motivation to grow professionally, intention and willingness to become a better teacher and the desire to make use of internal assessment to promote students' learning were some of the strong reasons for his excellent progress. Also, he spent a lot of time with the researcher discussing several aspects of CA and assessment scenario in his school. There was no doubt that he found the intervention programme to his liking and was happy to be a part of it.

Overview of Changes in Case 2 (State Board)

Though the two State Board teachers performed equally on all the sections of the CLAL instrument after the intervention, the percentage of improvement in their scores varied a lot from each other in the reading and writing sections. While the percentages of increase were around 35 and 66 for the third teacher, they were around 65 and 33 for the fourth

teacher for the aforesaid sections. In all other sections, the increase was same for both the teachers. They had problems in understanding the plan of assessment. Also, both of them had some difficulty in answering questions on assessment of grammar. It may be due to the fact that not much time was spent during the intervention programme on the assessment of grammatical ability. Even the improvement in feedback section can be attributed to the attention paid to it during the programme.

5.5.3 Case 3: Pre-, Post-Intervention CLAL Levels and Changes

Pre-intervention CLAL Levels

Teacher 5

The following figure displays the performance of *Teacher 5* on the CLAL instrument. The teacher scored 22 marks out of 38. So the level assigned to her was 'Limited'.
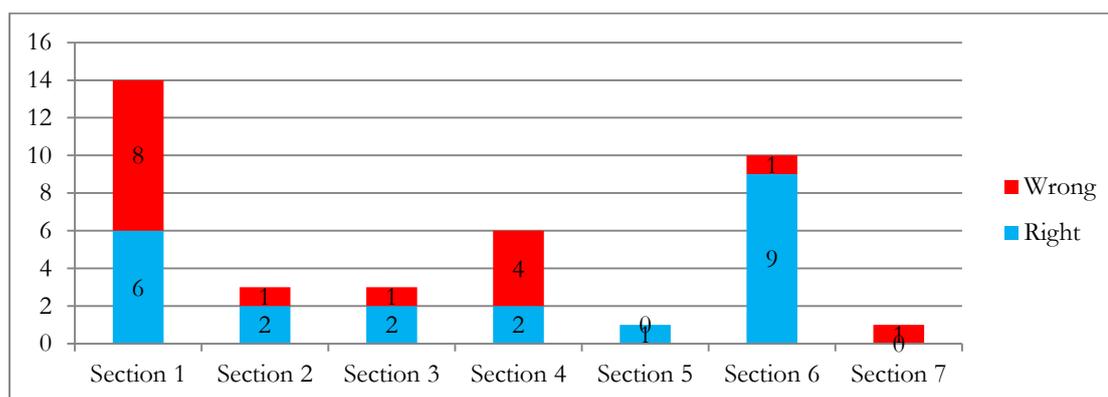
Figure 5.14: Performance of *Teacher 5* on CLAL instrument

She fell two marks short of the 'Average' level. Her score was remarkably high in section 6 which focused on classroom assessment and included questions on learner centered assessment, alternative assessment, washback, diagnostic assessment and authenticity. She also got the answer to the fifth question right. Apart from these, she got two out of three responses right for sections 2 and 3. In the rest of the sections, her performance was below average.

Teacher 6

With the score of 24, *Teacher 6* was the most successful among the six teachers who responded to the CLAL instrument. Apart from the only question in section 7, the teacher did not fail to get at least some of the answers right in all other sections.
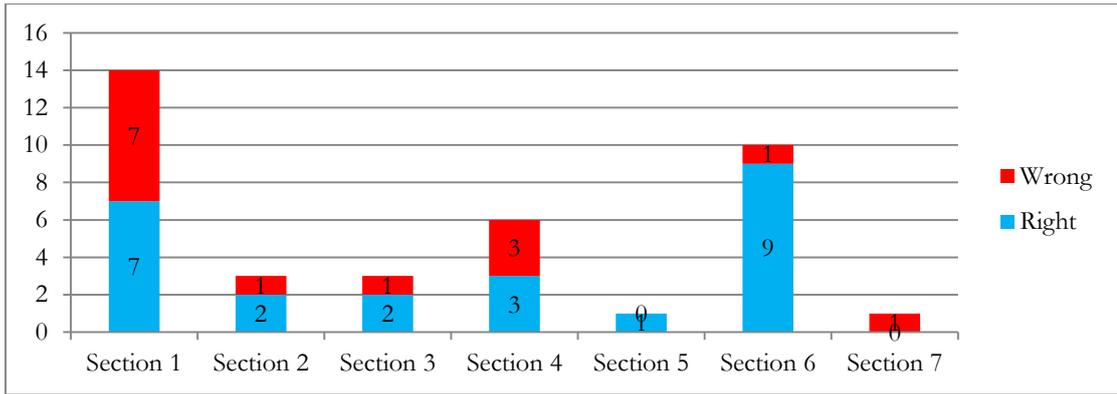
Figure 5.15: Performance of *Teacher 6* on CLAL Instrument

Like her colleague, she was right in her responses to questions on classroom assessment. In addition, she scored more than average in sections 2 and 3 which had questions on rater reliability in Writing and validity, authenticity and alternative assessment in the assessment of Speaking. She did not perform well in the first section which was on Reading and the fourth section which was on authentic assessment of grammatical ability.

Post-intervention CLAL Levels

Teacher 5 The performance of *Teacher 5* on the CLAL instrument is displayed below:
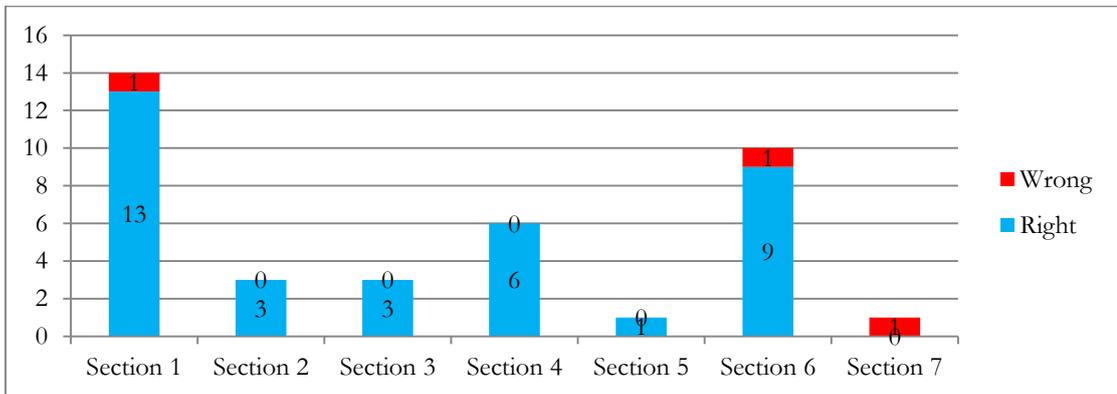
Figure 5.16: Performance of *Teacher 5* on CLAL instrument

Since she scored 35, she was considered to have 'Adequate' level of CLAL. Except in section 7 she got high scores in all other sections which included sections 2, 3, 4 and 5 in which she scored 100%. In the other two sections, she scored 90% and more than that. To be exact, she displayed excellent understanding of assessment of reading, writing, speaking and grammar and seemed to have a fair idea about the basic principles of

assessment, interpretation of scores, different types of assessment and classroom assessment.

Teacher 6

With a score of 37, *Teacher 6* gave the best performance on the CLAL instrument.
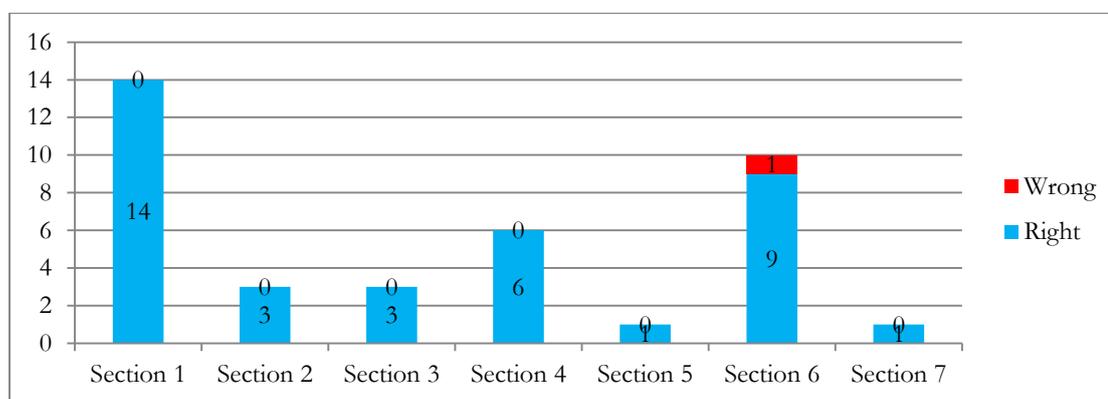
Figure 5.17: Performance of *Teacher 6* on CLAL instrument

The diagram above displays her performance in all the seven sections of the instrument. The only wrong response she gave was in section 6. Her performance proves that she had solid understanding of basic concepts pertaining to assessment of reading, writing, speaking, grammar, providing feedback and plan for designing classroom assessments. These also included application of basic principles of assessment, different types of assessment and task evaluation.

Change in Teachers' CLAL

Teacher 5

Like the other teachers, *Teacher 5* also displayed progress in CLAL through the TD programme. She recorded 34.21 % growth in her performance on the CLAL instrument. The percentage of improvement for individual sections is presented in the following diagram.

	Section 1	Section 2	Section 3	Section 4	Section 5	Section 6	Section 7
Pre-intervention	42.85	66.66	66.66	33.33	100	90	0
Post-intervention	92.85	100	100	100	100	90	0
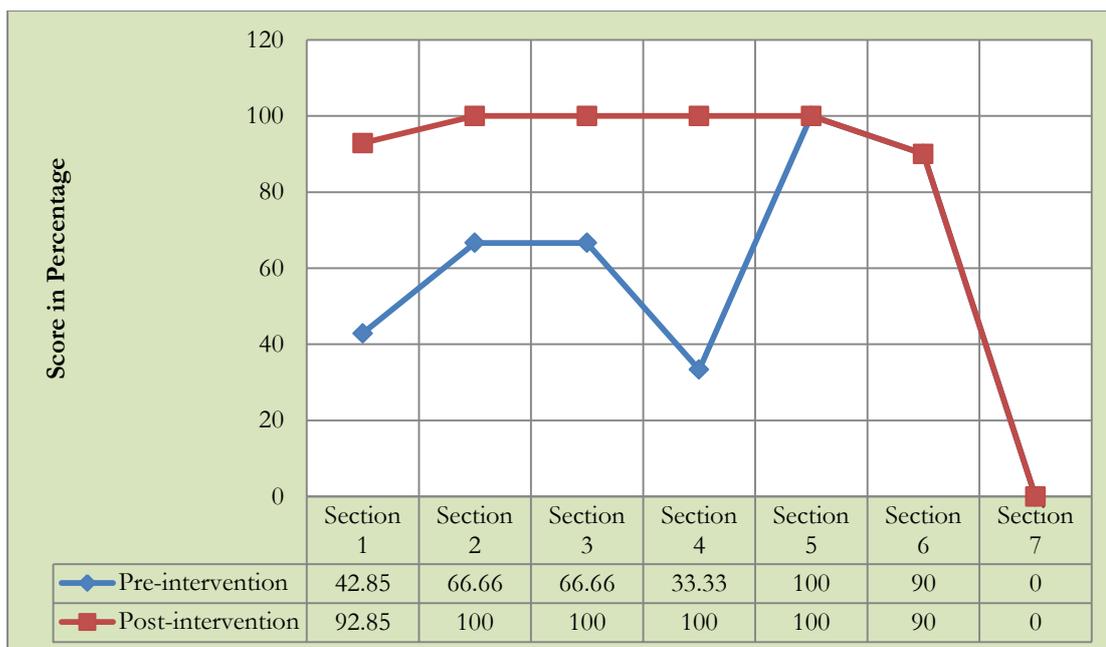
Figure 5.18: Pre- and post-intervention performance of *Teacher 5 on the* CLAL instrument

The teacher's scores for the sections on feedback, CA and plan of assessment did not change after the intervention, which can be interpreted mainly in two ways. The intervention either did not have any impact on the teacher's knowledge and beliefs in these areas or confirmed only some of the things related to the aforesaid areas the teacher knew and believed in. The teacher recorded healthy change in scores in sections on reading, writing, speaking and grammar with the development in grammar topping the sections. Though the teacher had a set of deep-rooted beliefs about assessment acquired through years of teaching experience, she still managed to score well. Moreover, her school did not give much importance to teacher-based assessments and the scope for using the knowledge gained through the intervention programme was completely up to her. The reasons for her achievement could include her proficiency in English, years of teaching experience, ability to connect the components of the intervention to classroom assessment situations and her desire to maintain herself as a competent teacher.

Teacher 6

This teacher got highest CLAL scores in both pre- and post-intervention surveys. She scored 37 in the post-intervention CLAL survey and recorded 34.21% of increase over her pre-intervention score. She could answer all the questions correctly except one in

126

section 6 during the post-intervention survey. The details are presented below in the diagram.
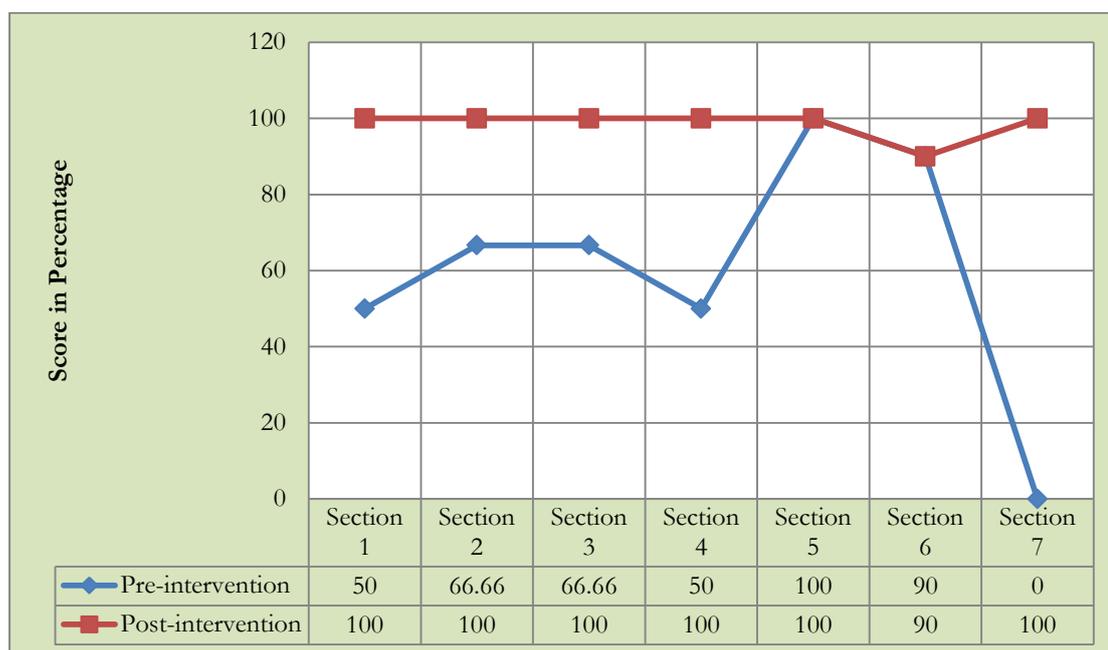
Figure 5.19: Pre- and post-intervention performance of *Teacher 6 on the* CLAL instrument

There was no difference between the pre- and post-intervention scores of the teacher in two sections of the instrument. It can be interpreted as either or negatively by relating the lack of change to the impact of the intervention programme on the teacher's knowledge and beliefs. While sections on reading and grammar recorded 50% increase and the ones on writing and speaking, around 33%, the teacher answered the question on the plan of assessment correctly. Like her colleague, she was not much interested in the TD programme at the beginning and had little expectation from it. Even then, she seemed to have understood most of the basic concepts of assessment presented during the intervention programme. At least, her performance on the CLAL survey instrument indicates so. Her language ability, focus and a constant desire to establish herself as a competent teacher could be some of the reasons why she made this progress.

Overview of Case 3 (ICSE)

The ICSE teachers showed a lot of similarity in the way their scores changed after the TD programme. Along with the same overall increase of 34.21%, their scores matched in four individual sections— writing, speaking, feedback and CA. While no changes were

observed in the sections on feedback and CA, the writing and speaking sections moved up by around 33%. The high scores in feedback and CA before the intervention can be interpreted, though without much certainty. But the continuation of the same scores may invite various interpretations as already mentioned in the sections talking about changes found in the individual teachers under Case 3. Even the 100% difference in the last section cannot be interpreted conclusively.

5.5.4 Cross-Case Analysis

All the participant-teachers showed a lot of improvement in their CLAL after undergoing the TD programme in CA. Since the teachers were paired under three cases with each case comprising two teachers from a different boards of education, the average improvement for each case is also highlighted in the following diagram.
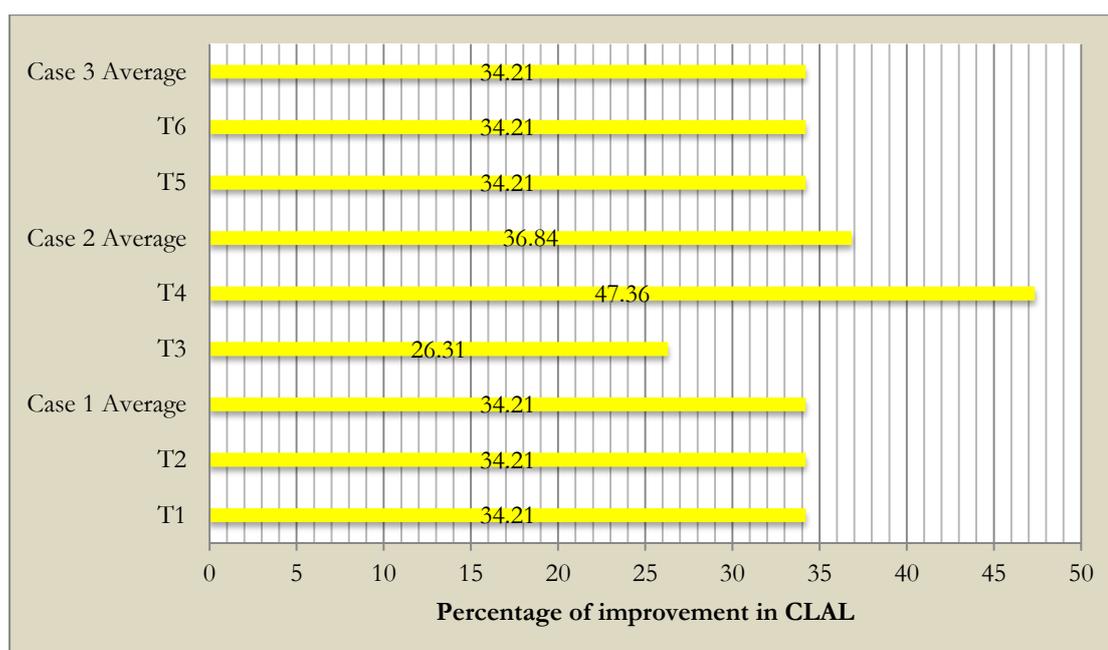
Figure 5.20: Development in CLAL level of the teachers

As highlighted in the above diagram, there is not much difference in the percentage of improvement achieved by the teachers in terms of scores on the CLAL instrument across cases. But their post-intervention scores were much better than their respective pre-intervention scores. The average scores of improvement for Case 1, Case 2 and Case 3 were 34.21%, 36.84% and 34.21% respectively. Apart from *Teacher 4*, who registered an improvement of 47.36%, all others exhibited similar percentage of achievement. These

scores imply that the TD programme was effective in developing the teachers' CLAL, and all the six teachers responded to it in a positive manner irrespective of the educational boards under which their schools functioned. However, the changes in individual sections sometimes varied from one case to another. The most obvious of such changes can be found in the feedback section in which the first two cases displayed 100% increase, whereas Case 3 had a 100% score even before the intervention. Another noticeable thing is the teachers' scores in the section on the plan of assessment. They had problems either in understanding the concept or the way the question has been constructed. Though it was also observed that a need-based TD programme can have positive impact on teachers' professional knowledge-base, differences among teachers in each case indicated that teachers can have individual responses depending on several contextual and personal factors.

5.6 Conclusion

This chapter has presented the data about the state-wide survey of secondary school English teachers' CLAL, the development and conduct of the TD programme and its impact on the participating teachers' CLAL. The impact of the intervention on their beliefs about assessment practices, ability to prepare appropriate CA tasks, develop relevant assessment criteria and provide feedback is presented in the next chapter. In addition, the relationship between the teachers' CLAL and, their ability to carry out CAs and beliefs about CA is also discussed there.

CHAPTER 6

EXAMINATION OF THE RELATIONSHIP BETWEEN TEACHERS' CLASSROOM LANGUAGE ASSESSMENT LITERACY AND ABILITY TO CARRY OUT CLASSROOM ASSESSMENTS

6.1 Introduction

This chapter is devoted to the analysis and interpretation of the data related to the relationship between teachers' CLAL and assessment practices. The chapter has three parts. The first part focuses on the change in teachers' beliefs about assessment, the second part, on the changes in teachers' ability to design assessment tasks and criteria and offer feedback, and the last part discusses the relationship between teachers' CLAL and their assessment practices.

6.2 Changes in Beliefs about Assessment

A questionnaire (see CHAPTER 4, section 4.10.2) was used to capture the beliefs of six teachers, i. e. the ones chosen to participate in the intervention programme, about classroom assessment of language ability. The data obtained through the questionnaire are discussed below under three loosely divided clusters: Classroom Assessment (CA), Principles of Assessment (PA) and Teachers' Assessment Ability (TAA) for each case. For analyzing the responses of teachers to the belief questionnaire, the options- *SA, A, D, SD* and *DN-* were assigned numerical values. Thus, *SA, A, DN, D* and *SD* were coded as 2, 1, 0, -1 and -2 respectively. The analysis takes into account the numerical data and data collected through interviews, classroom observations, etc. to interpret the results. For each case, at first, the beliefs of individual teachers before and after the TD programme are presented along with the changes observed in each teacher. An overview of changes in each case follows the section on changes in individual teachers. A cross-case analysis is taken up in the last part of the section.

6.2.1 Case 1: Pre-, Post-Intervention Beliefs about Assessment and Changes

Pre-intervention Beliefs

Teacher 1

The beliefs of *Teacher 1* about assessment in the three sections CA, PA and TAA, as shown in Figure 6.1, suggest that the teacher had strong positive beliefs about the aforesaid areas. The overall score 1.33 (CA = 1.6, PA = 0.83 and TAA = 1.58) supports this conclusion to a great extent. The mean score for the teacher in CA was 1.6. It means, at least theoretically if not in practice, the teacher was aware of certain basic things about CA, such as having regular classroom assessments rather than one final examination, getting information about students' language ability in more than one way, using information from assessment for teaching purposes, involving students in the process of assessment and integrating assessment with teaching. However, the mean score for PA was less than half of that of CA. It indicates that the teacher had less positive and sometimes negative beliefs about some of the principles of assessment. The teacher disagreed with the view that only the language skills which are taught in the classroom should be assessed. She also disagreed that listening and speaking can be assessed in the classroom without a tape recorder and an audio player. The response to the statement that "[t]he test score should show the real ability of a student to use language" suggests that the teacher was either not well informed about the concept of language ability or not sure about the concept of validity. But she was sure that students should not memorise and reproduce information from the text book during classroom assessment, and that reading comprehension should be assessed using unseen texts and grammar and vocabulary should be assessed in communicative context.

The teacher strongly agreed to most of the statements covered under the section TAA as evident from the mean score 1.58 for the section. Except for the last statement in the questionnaire concerning the consideration of factors such as class, caste, gender and religion while designing tasks, the teacher responded to all the statements positively. While she agreed with eight statements strongly, she showed just agreement with statements focusing on allowing extra time to slow writers and encouraging students to learn as one of the purposes of assessment. More importantly, the teacher strongly believed that teachers need training in language assessment and should have knowledge about language proficiency and basic principles of language assessment.

Teacher 2

When compared to the first teacher the second teacher, who belonged to the same school as the first teacher, showed stronger beliefs about the aspects of assessment covered in the belief questionnaire. The overall average score of 1.73, which was also the highest score among all the teachers, suggests that the teacher believed very strongly in the statements included in the questionnaire. For CA, he had a mean score of 1.8. In the section he very strongly agreed with formative assessment, classroom assessment as an aid to the teacher, involving students in the process of assessment and integrated assessment. Similarly, he got a mean score of 1.83 for PA. He seemed to be convinced that test scores should reflect students' language ability, not their ability to memorise and reproduce information. He believed in the employment of texts outside the syllabus for designing assessment tasks and providing context while assessing grammar and vocabulary.

Though the mean score for the next section, i.e., TAA, was slightly lower than those for the previous sections, yet the score 1.58 suggests strong agreement. For statements 5, 6, 13, 15, 20, 21, 22, 23 and 24, he chose to agree strongly. These included statements focusing on the necessity of training in language assessment for teachers, understanding of language proficiency, informing students about their performance, positive washback of assessment, individual feedback to students, validity of assessment tasks, awareness about basic principles of assessment and analysis and reporting of assessment results. However, much like the first teacher the teacher disagreed with consideration of factors like class, caste, etc. while constructing assessment tasks.

Post-intervention Beliefs

Teacher 1

The overall average score for *Teacher 1* was 1.82. It is obvious that the teacher had strong beliefs about the aspects of assessment covered in the questionnaire. The teacher got mean scores 1.8, 1.83 and 1.83 for CA, PA and TAA respectively. She did not disagree with any of the statements. Moreover, she strongly agreed with 19 out of 23 statements mentioned in the questionnaire. In the section on classroom assessment, the teacher chose not to agree strongly with the idea of involving students in the process of assessment. She maintained the same attitude for the statement numbers 14, 17 and 27.

Teacher 2

The second teacher, i.e., *Teacher 2,* also had a very high overall average score like that of his colleague. In fact, his overall average score- 1.88 was the highest among all the six teachers. He got a perfect mean score of 2 for his responses in the section CA. Except for the statement that "[l]istening and speaking can be assessed by the teacher in the classroom even without a tape recorder and an audio player", he strongly agreed with all other statements in the section PA. For the third section his mean score was 1.83. He showed strong beliefs about statements related to teachers' assessment ability. Only on two occasions i.e., for statements 17 and 27, he chose to agree rather than strongly agree.

Changes in Beliefs

Teacher 1

The overall mean score for the first teacher on the pre-intervention questionnaire was 1.33, whereas it was 1.82 on the same questionnaire administered after the intervention. The following diagram presents the comparison.

	Mean Scores			
	Classroom Assessment (CA)	Principles of Assessment (PA)	Teachers' Assessment Ability (TAA)	Overall Average Score
Pre-intervention	1.6	0.83	1.58	1.33
Post-intervention	1.8	1.83	1.83	1.82

Figure 6.1: Pre- and post-intervention teacher beliefs scores of *Teacher 1*

The increase in the overall average score indicates that the teacher's beliefs about classroom assessment became more positive and stronger after the intervention. Even for the individual sections-CA, PA and TAA, there was increase in the mean scores. The gap was very high for the section PA. The pre-intervention mean score for this section was 0.83, and it leapt to 1.83 in the post-intervention survey. This change could have been due to many reasons. It could be that the teacher was not aware of the five basic principles of assessment, which he was exposed to during the TD programme. It is also possible that the teacher might have found the principles immediately useful and

convincing. It is essential to notice that some of the negative beliefs became positive after the intervention.

Teacher 2

The second teacher showed even stronger positive beliefs about classroom assessment than his colleague in his responses to the questionnaire before and after the intervention. The details of a comparison are highlighted below in the diagram:

	Mean Scores			
	Classroom Assessment	Principles of Assessment	Teachers' Assessment Ability	Overall Average Score
Pre-intervention	1.8	1.83	1.58	1.73
Post-intervention	2	1.83	1.83	1.88

Figure 6.2: Pre- and post-intervention teacher beliefs scores of *Teacher 2*

His pre- and post-intervention overall mean scores were 1.73 and 1.88 respectively. Technically, even before the intervention, he had some awareness about aspects of assessment. The intervention helped to reinforce these beliefs and make them stronger. That is why the only noticeable change was in the mean score for the section TAA which went up by 0.25. The teacher strongly agreed with quite a few statements in all the three sections of the questionnaire and tried to establish that he was familiar with concepts like language ability, formative assessment, learner-oriented assessment, positive washback, etc.

Overview of Changes in Case 1

The changes in the beliefs of the CBSE (Case 1) teachers had a few similarities. The positive beliefs of the teachers after the TD programme yielded the similar scores. However, changes recorded in the individual sections in the questionnaire varied from one teacher to another. The most striking one was the difference between their scores in the section PA. The significant increase in the first teacher's score in the section can be contrasted with the absence of any change in the second teacher's score for the same section. This difference may not have anything to do with beliefs as the second teacher was little aware of language assessment as evident from his pre-intervention assessment

tasks and classroom teaching. On the contrary, the first teacher had some idea about language assessment and she prepared better tasks than the first teacher. In the two other sections, i. e., CA and TAA, both the teachers did not get significantly different scores. Their high scores in CA are, once again, quite puzzling because all the statements comprising this section are related to the importance, benefits and principles of CA and the teachers were found lacking adequate knowledge about CA. It is possible that the teachers might have heard about these things since the school follows CCE and the CCE documents discuss the above mentioned things. But the scores did not change much after the intervention which might indicate that the intervention either contributed to the strengthening of or had little impact on what they claimed to believe in.

6.2.2 Case 2: Pre-, Post-Intervention Beliefs about Assessment and Changes

Pre-intervention Beliefs

Teacher 3

The third teacher, who worked in a state-board school, showed positive beliefs about assessment. Though her overall average score, i.e., 1.18, was lower than those of the pair of teachers from the CBSE school, still, she was on the same continuum of "agree— strongly-agree". For the first section, her mean score was 1.8 which was on a par with the CBSE teachers. She strongly agreed with four out of five statements in the section. These statements contained details about preference of formative over summative assessments, classroom assessment as an instructional tool, student involvement in classroom assessment and integrating classroom assessment with teaching. In contrast to her mean score in the first section, i. e., 1.8, her mean score for PA suggested very little agreement with the statements. Though she strongly agreed that test scores should reflect students' ability to use language, she disagreed that only those skills which are taught in the classroom should be assessed and believed that listening and speaking can be assessed without a tape recorder or an audio player. Her beliefs about PA also included agreement with skills assessment, as opposed to memory tests, and assessment of grammar and vocabulary in appropriate context.

For the section covered under TAA, the mean score of teacher responses was 1.25. The teacher strongly disagreed only with the last statement. For all other statements, the teacher showed positive beliefs. She strongly agreed that teachers should have training in language assessment, understanding of the concept of language proficiency, positive

washback of assessments, understanding of assessment objectives and basic principles of assessment and that it is necessary for teachers to be able to report assessment results effectively. Although she did not agree strongly, nevertheless she agreed that students should be informed about the assessment criteria in advance, they should be provided with feedback about their performance, the questions should be of appropriate difficulty level and slow writers should be allowed a little extra time.

Teacher 4

With the overall average score of 0.8 *Teacher 4* got the lowest score among all the six teachers. Though the score falls on the "don't know—agree" continuum and shows overall positive attitude of the teacher towards the aspects of assessment covered in the questionnaire, yet on many occasions the teacher disagreed with the statements. In the first section, his mean score was 0.6. He agreed with statements related to adopting formative assessments in place of summative ones, accepting classroom assessment as an asset to teaching, involving students in the process of assessment and assessing students while teaching. The only statement with which he disagreed was the one which emphasised giving students more than one chance to display their ability to use language. This is self-contradictory because he agreed to the first statement in the section that there should be regular classroom assessments rather than one final examination.

In the next section which focused on PA, the mean score was 0.66. So the level of agreement was as positive as it was for the first section. He did not believe in the statements that "[a]n assessment should not assess students' ability to memorise and reproduce information from their textbook" and "[o]nly those language skills which are taught should be assessed". On the other hand, he agreed with two of the statements and strongly agreed with two more. He seemed to be very convinced that listening and speaking can be assessed with an audio player and grammar and vocabulary should be assessed in communicative context.

For the statements under the section TAA, the teacher got a better mean score than what he did for the previous two sections. The score 1.16 may mean that the teacher showed more than mere agreement with the statements. Except for the last statement focusing on caste, class, gender, etc., as factors to be considered during the construction of any assessment, with which he strongly disagreed, on all other occasions he either agreed or strongly agreed with the statements. The most notable one was his strong agreement

with the statement that teachers need training in language assessment to function effectively.

Post-intervention Beliefs

Teacher 3

This teacher showed positive beliefs about assessment in her responses to the teacher belief questionnaire. Her overall average score was 1.68 and her scores for CA, PA and TAA were 1.8, 1.66 and 1.58 respectively. She strongly agreed with all the statements clubbed under CA but preferred just to agree that students should be given multiple chances for exhibiting their language ability. For the next section, she agreed with statements focusing on validity of test scores, skills-based assessments, assessment of reading comprehension using unseen texts and providing context in tasks meant for assessing students' grammar and vocabulary abilities. For the section TAA, the teacher strongly agreed with seven statements and agreed with rest of the five statements.

Teacher 4

The fourth teacher, who was a colleague of *Teacher 3,* got a similar overall average score like that of his colleague. His overall score of 1.67 which was on the second half of the "agree—strongly agree" continuum showed his positive beliefs about assessment. Out of the five statements in the section CA, he strongly agreed with three of them and had a mean score of 1.6. For the next section, which focused on principles of assessment, his mean score was 1.83. In this section, he showed strong agreement with all the statements except the statement (no. 10), which referred to the use of unseen texts for assessing reading comprehension skills. For the section TAA, he got 1.58 which was his lowest mean score. The ratio of "strongly agree—agree" responses for this section was 7:5. It suggests his positive and strong beliefs about the importance of teachers' assessment ability.

Changes in Beliefs

Teacher 3

The difference between the overall mean scores obtained by *Teacher 3* before and after the intervention was 0.5. More details about the change in beliefs are presented in the following diagram.

	Mean Scores			
	Classroom Assessment	Principles of Assessment	Teachers' Assessment Ability	Overall Average Score
Pre-intervention	1.8	0.5	1.25	1.18
Post-intervention	1.8	1.66	1.58	1.68

Figure 6.3: Pre- and post-intervention teacher beliefs scores of *Teacher 3*

Her pre-intervention overall mean score was 1.18, which indicates that she had positive beliefs about classroom assessment even before the intervention. There was disparity among the three pre-intervention mean scores- 1.8, 0.5 and 1.25, for the three sections- CA, PA and TAA respectively. The teacher got negative scores twice for the section PA and once for the section TAA. In contrast to these pre-intervention scores, she scored 1.8, 1.66 and 1.58 for CA, PA and TAA respectively for her response to the questionnaire administered after the intervention. Moreover, there was not even a single disagreement with any of the statements in her post-intervention response.

Teacher 4

The fourth teacher scored overall mean scores 0.8 and 1.67 for his responses to the beliefs questionnaire before and after the intervention respectively. The following diagram contains the details.

	Mean Scores			
	Classroom Assessment	Principles of Assessment	Teachers' Assessment Ability	Overall Average Score
Pre-intervention	0.6	0.66	1.16	0.80
Post-intervention	1.6	1.83	1.58	1.67

Figure 6.4: Pre- and post-intervention teacher beliefs scores of *Teacher 4*

Though his pre-intervention score is a positive one and falls on the continuum "don't know-agree", some of his responses were negative— one in the section CA, two in PA and one in TAA. It is important to mention that some of his responses were self-

contradictory which might mean that his original thinking was challenged after he underwent the TD programme. The score 0.8 was also the lowest among all the teachers. But the post-intervention scores for individual sections suggest that the teacher developed positive beliefs through the intervention programme. He responded positively to statements related to giving students more than one chance to display their language ability, thinking about issues like caste and gender while constructing assessment, etc. in his post-intervention response. In fact, these were the statements he did not agree with before the intervention. These changes may not indicate complete internalization of concepts related to CA. They, rather, suggest recognition of the concepts.

Overview of Changes in Case 2

Though there were some differences in the pre-intervention scores of the two State Board teachers, the post-intervention mean scores were almost the same. The beliefs they shared after the intervention were identical. Except for the section on CA, in which the changes in beliefs varied greatly, the changes in the other two sections were identical. The third teacher reported that she did not like the change in assessment policy and felt lost while trying to know the utility of FAs. Thus, her high pre-intervention score in the section CA indicated her lack of understanding of the statements. In case of the fourth teacher, the concept of CA/FAs perhaps started making sense after the intervention. The change in the teachers' beliefs in the other two sections, i. e., PA and TAA, were in line with the school assessment culture and the board assessment policy. But it is difficult to establish a uniform connection between the TD programme and the changes in beliefs.

6.2.3 Case 3: Pre-, Post-Intervention Beliefs about Assessment and Changes

Pre-intervention Beliefs

Teacher 5

The first of the ICSE teachers got the second lowest overall average score among all the six teachers who participated in the intervention. Her mean score for CA was 0.16. On the one hand, she disagreed with the concept of formative assessment; on the other hand, she agreed that students should be given multiple chances to show their language ability and also, strongly agreed that assessment can be integrated with teaching. She strongly disagreed with involving students in the process of assessment.

The teacher got a mean score of 1.33 for TA. As shown in the Teacher Beliefs Table, she did not believe in assessing only language skills that taught. She strongly agreed that test scores should reflect students' language ability, they should not be asked to memorise and reproduce information during assessments, reading comprehension should be assessed using unseen texts and assessing oral abilities without a tape recorder or an audio player.

The teacher got a positive mean score in the section TAA. She preferred not to agree or disagree with the idea of informing students in advance about the assessment criteria and having assessments that encourage learning. She strongly disagreed, like *Teacher 3* and *Teacher 4*, that caste, class and gender, etc., should be taken into account while designing assessments. Nevertheless, she strongly agreed that teachers need training in language assessment; they should inform the stakeholders about students' performance, consider the difficulty level of the tasks, provide individual feedback to students, understand objectives of the assessment and have an understanding of basic principles of assessment.

Teacher 6

The sixth teacher got an overall average of 0.97 which indicates her positive beliefs about language assessment, especially about those aspects included in the questionnaire. Her mean score for CA was 0.33 which was the second lowest among the scores by all the six teachers for the section. Like her colleague, i.e., *Teacher 5*, she did not believe in having formative assessment in place of summative ones and involving students in the process of assessment. She almost self-contradicted herself by agreeing that students should be given more than one chance to exhibit their language ability and that assessment should be integrated with teaching.

For the statements under the section PA, she got 1.33 which shows four times stronger and more positive belief than what she got for CA. She strongly agreed that test scores should truly reflect students' language ability, classroom assessment should not focus on testing students' memory, reading comprehension skills should be assessed using unseen texts and oral skills can be assessed without using a tape recorder or an audio player.

The teacher got a mean score of 1.25 for the section TAA. Like all other teachers she disagreed with the last statement concerning caste, class and gender, etc. as factors to be considered while designing assessment tasks. To all other eleven statements, she

responded positively. She showed strong belief in the necessity of training in language assessment for teachers, giving students feedback about their performance, understanding objectives of lessons and assessments and the basic principles of assessment.

Post-intervention Beliefs

Teacher 5

The fifth teacher who worked in ICSE school got an overall average score of 1.41. In the section CA, she got the mean score of 1. She agreed with statements involving formative assessment, giving multiple chances to students for showing their language ability and classroom assessment as an asset to teaching. However, she took a neutral stand about students' involvement in the process of assessment. In the section focusing on principles of assessment, she strongly agreed with four statements and just agreed with two statements. She preferred not to strongly agree with assessment of only those language skills that are taught in the classroom and using enough context for the assessment of grammar and vocabulary. For the statements under the category TAA, the teacher's mean score was 1.58. Once again the ratio between strongly-agreed and agreed statements was 7:5. It indicates strong positive beliefs of the teacher.

Teacher 6

This teacher got the overall average score of 1.44 in her responses to the belief questionnaire after the intervention. Her mean scores for CA, PA and TAA were 1, 1.83 and 1.5 respectively. The teacher responded to the statements under CA exactly as her colleague. She chose to be neutral about involving students in the process of assessment but strongly agreed that assessment can be integrated with teaching. She strongly agreed to all the statements under the category PA except the one focusing on providing enough contexts to students during the assessment of grammar and vocabulary. As evident from her coded responses mentioned in the Post-Intervention Teacher Beliefs Table 1, she strongly agreed with six of the total twelve statements and merely agreed with the rest of the six.

Changes in Beliefs

Teacher 5

The fifth teacher had the second lowest overall pre-intervention mean score among all the teachers. In contrast to her pre-intervention score of 0.85, she got 1.41 for her post-intervention responses to the beliefs questionnaire. The following diagram presents her scores in the individual sections.

	Mean Scores			
	Classroom Assessment	**Principles of Assessment**	**Teachers' Assessment Ability**	**Overall Average Score**
Pre-intervention	0.16	1.33	1.08	0.85
Post-intervention	1	1.66	1.58	1.41

Figure 6.5: Pre- and post-intervention teacher beliefs scores of *Teacher 5*

She disagreed with five statements among which there were two strong disagreements, before the intervention. Her pre-intervention mean score for the section CA was 0.16—the lowest score for any section by any teacher. The disparity between the scores obtained for CA and the two other sections existed even for her post-intervention responses. Her post-intervention mean score for CA was 1, whereas, for the sections PA and TAA, the scores were 1.66 and 1.58 respectively. She showed change in beliefs by agreeing with statements related to formative assessment and considering issues like gender and caste while constructing assessments in her post-intervention responses. This was conspicuous because she disagreed with all these statements before the intervention. It was, however, surprising that she chose to remain neutral about students' involvement in the process of assessment.

Teacher 6

With overall mean scores 0.97 and 1.44 for pre- and post-intervention responses respectively, some change in beliefs was observed in *Teacher 6*. A comparison between pre- and post-intervention scores in individual sections of the questionnaire is presented below.

Mean Scores			
	Classroom Assessment	Principles of Assessment	Teachers' Assessment Ability	Overall Average Score
Pre-intervention	0.33	1.33	1.25	0.97
Post-intervention	1	1.83	1.5	1.44

Figure 6.6: Pre- and post-intervention teacher beliefs scores of *Teacher 6*

The change was particularly very obvious for the section CA in which she got 0.33 as her pre-intervention score and 1 as her post-intervention score. Though she responded negatively to four of the statements in the questionnaire before the intervention, she disagreed with two statements— one proposing formative assessment over summative and the other, involving students in the process of assessment— in the CA section. Also, some of her pre-intervention responses were self-contradictory and thus, indicated her lack of conviction about different aspects of classroom assessment. She did not disagree with any of the statements during the post-intervention survey which suggests some amount of change in her belief system which could be due her improved CLAL level.

Overview of Changes in Case 3

The ICSE (Case 3) teachers had more common beliefs about assessment than the CBSE and the State Board teachers. In all three sections, the changes for both the teachers were unidirectional. They did not agree with many of the statements before the intervention, which might have been a direct impact of the school assessment system and to some extent, also the ICSE assessment policy. Their beliefs did not contradict what they reported practicing and liking. The changes in beliefs, however, did not reflect their existing doubts about different aspects of CA and the need for enhancing teachers' assessment ability. The reported changes, thus, were far less conclusive than those found in the other two cases. This belief gets further strengthened by the interview and classroom observation data.

6.2.4 Cross-case Analysis of Changes

A comparison among the overall pre- and post-intervention average scores of the three cases along with the teachers' average scores for each of the three sections of the beliefs

questionnaire and the same for each case are presented in the following table (Table 5.6). Though Case 1 recorded the highest overall case average pre- and post-intervention scores, the percentage of increase in the overall case average was highest for Case 2. Case 2 exhibited an increase of 40.71%, whereas the percentage was 35.91 for Case 3, 17.29 for Case 1 and the average gain across all three cases was 30.9%. The increase in scores for the State Board teachers can be interpreted as an impact of the exposure to new knowledge about assessment provided through the TD programme. Among all the three pairs, they had perhaps the best chance of utilizing and testing the new knowledge skills in their school. The princess did not wish to continue in that place because the monster.

Table 6.1: Teachers' Beliefs about Assessment: A Cross-case Comparison of Scores

Case	Teacher	Classroom Assessment				Principles of Assessment				Teachers' Assessment Ability				Overall Average Score		Overall Case Average	
		Teacher Average		Case Average		Teacher Average		Case Average		Teacher Average		Case Average					
		Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post	Pre	Post
1	1	1.6	1.8	1.7	1.9	0.83	1.83	1.33	1.83	1.58	1.83	1.58	1.83	1.33	1.82	1.53	1.85
	2	1.8	2			1.83	1.83			1.58	1.83			1.73	1.88		
2	3	1.8	1.8	1.2	1.7	0.5	1.66	0.58	1.74	1.25	1.58	1.2	1.58	1.18	1.68	0.99	1.67
	4	0.6	1.6			0.66	1.83			1.16	1.58			0.8	1.67		
3	5	0.16	1	0.24	1	1.33	1.66	1.33	1.74	1.08	1.58	1.16	1.54	0.85	1.41	0.91	1.42
	6	0.33	1			1.33	1.83			1.25	1.5			0.97	1.44		
Average →				1.04	1.53			1.08	1.77			1.31	1.65			1.14	1.65

Apart from the section on CA, in which Case 3 teachers had a higher percentage of increase, Case 2 teachers showed stronger change in beliefs in the sections, PA and TAA.

Among the overall average scores obtained by the three cases in the three individual sections — CA, PA and TAA, the highest average improvement was recorded for CA. The pre-intervention score was highest for TAA. This could be interpreted as teachers' awareness about the importance of assessment ability for language teachers.

A few striking differences among cases in terms of scores were observed. The most obvious one was the difference between pre- and post-intervention scores of Case 3 teachers on the one hand and the CBSE and State Board teachers on the other hand, for the CA section. The ICSE teachers did not show much positive belief in formative assessment covered under the aforesaid section. The lack of policy-level emphasis on CA could be a reason behind that.

Another surprising difference was the pre-intervention score of State Board (Case 2) teachers in the PA section. It was less than half of what the other cases got. However, the Case 2 teachers had a highly improved post-intervention score in that section, on par with the other two pairs. The low pre-intervention and high post-intervention scores in PA of Case 2 teachers, to some extent, indicated positive change in their beliefs about assessment. It could be safely assumed that the intervention must have contributed to this change.

6.3 Changes in the Ability to Design Assessment Tasks and Criteria and Offer Feedback

As per the plan, information about the ability of the teachers to construct appropriate assessment tasks was collected by the researcher. As it was not feasible to get real classroom assessment tasks from each teacher due to practical constraints, it was necessary to make all the six teachers design tasks for assessing the four major language skills along with vocabulary and grammar outside their classroom context. The target group of students was their own students. Some teachers designed the tasks individually but some others discussed the tasks with the other teacher while designing them. They were given rough parameters which included a list of components (aim of the task, assessment criteria, types of questions and feedback) to be included in each task. They were, however, not ready to design the entire task with the questions/activities for all the skills and components. So the researcher agreed to accept the information provided by

each teacher about task design and interviewed them whenever any required information was found missing from their tasks.

The assessment tasks were rated by three researchers including the researcher. All of them had training in language assessment. Both the other raters were made familiar with the rating scale and what to look for in the tasks. They were also encouraged to ask questions of the researcher about the tasks if they found anything missing in them. This process was necessary because none of the teachers provided the complete description of tasks. It was only during the interviews taken after they constructed the tasks that they provided further details about certain tasks.

After obtaining three scores for each characteristic and each task (for each skill and component), the average score was calculated for them. In this step, an average score is arrived at for each teacher. The same process is repeated for post-intervention tasks too. The scores and the calculations for the pre-intervention teacher-made tasks are presented in a tabular form in the next few sections.

Four levels were created and each level was decided on the basis of overall scores assigned to teachers. The details about all the levels are presented below:

Table 6.2: Level descriptors for teachers' ability to design assessment tasks

Level/ Scores (from-to)	Description
Excellent 4-5	Able to: • create learning- and learner-friendly classroom assessment tasks effectively • follow the basic principles of assessment across tasks • cater to the learning needs of the target group to a great extent
Good 3-4	Able to: • create learning- and learner-friendly classroom assessment tasks with very few problems • follow the basic principles of assessment on most occasions • cater to the learning needs of the target group to some extent
Average 2-3	Able to: • create classroom assessment tasks but with a lot of difficulty • follow only one or two principles of assessment on some occasions • cater to some of the learning needs of the target group on some occasions
Below Average 1-2	Needs to: • learn how to create classroom assessment tasks • understand the basic principles of assessment • understand the learning needs of the target group

In the tables, in which the scores for tasks designed by each teacher is presented, the letters 'L', 'S', 'R', 'W', 'V' and 'G' stand for Listening, Speaking, Reading, Writing, Grammar and Vocabulary respectively. It needs to be mentioned that in the tables presented for five other teachers, these letters have been removed and only task numbers stand for the skills in the order as shown in the above table. So 'Task 1' represents Listening Skills, 'Task 2' Speaking Skills and so on. The ten task characteristics can be found in APPENDIX F.

After analysing and presenting the changes in each teacher and then, each case, a cross-case analysis is taken up at the end of this section.

6.3.1 Case 1

Pre-intervention Ability

Teacher 1

The following table shows the levels of tasks prepared by *Teacher 1*. It also provides information about how the tasks fared against individual task characteristic and how each task was rated by the three raters.

Table 6.3: Pre-intervention task levels for *Teacher 1*

Task characteristic	Levels assigned to tasks by 3 different scorers						Average level for each characteristic
	Task 1 (L)	Task 2 (S)	Task 3 (R)	Task 4 (W)	Task 5 (G)	Task 6 (V)	
1	3/2/3	2/2/2	2/1/1	1/1/1	2/1/1	2/2/2	1.72
2	2/1/1	1/1/1	2/1/2	2/1/1	1/1/1	1/1/1	1.22
3	2/2/2	2/3/2	1/2/1	2/2/2	3/1/2	2/2/2	1.94
4	2/1/1	1/2/2	3/2/2	1/2/1	1/1/1	1/2/1	1.5
5	2/1/1	1/1/1	2/2/2	1/2/2	1/1/1	1/2/1	1.38
6	3/2/2	1/1/1	3/2/2	1/2/1	1/1/1	1/2/1	1.55
7	3/2/2	1/1/1	2/2/2	1/2/2	1/1/1	1/2/1	1.55
8	4/3/3	1/1/1	1/2/1	2/1/1	3/2/2	3/3/3	2.05
9	1/1/1	2/1/2	2/1/1	1/2/1	2/1/2	2/3/3	1.61
10	1/1/1	1/1/1	1/1/1	1/2/1	1/1/1	1/1/1	1.05
Average level of tasks →	1.86	1.36	1.66	1.43	1.33	1.7	
Overall level = $\dfrac{\text{Average level of tasks}}{6} = 1.55$							

As evident from the overall level of the tasks, i. e., 1.55 out of 5, the tasks designed by *Teacher 1* were rated quite low. Though the scores 1.36 and 1.33 indicate that the tasks meant for assessing Speaking and Grammar were not as good as the other tasks, yet the fact that none of the others were rated higher than 2 also proves that the tasks were not at all up to the mark. From another angle, i. e., the average level for each characteristic, the teacher's score in providing feedback was the lowest. It was followed by low scores in task validity (characteristic 2) and diagnostic information (characteristic 5). Apart from reliability of scoring, the scores for other characteristics remained below 2, i. e. average level.

The table presented below contains information about how the pre-intervention assessment tasks were rated and the levels assigned to them.

Table 6.4: Pre-intervention task levels for *Teacher 2*

Task characteristic	Levels assigned to tasks by 3 different scorers						Average level for each characteristic
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	3/2/3	2/2/2	3/2/2	3/3/1	3/2/3	1/1/1	2.16
2	1/1/1	3/1/2	1/1/1	3/3/3	1/1/1	1/1/1	1.5
3	3/2/2	3/3/3	3/2/3	3/3/3	1/1/1	1/1/1	2.16
4	1/1/1	1/1/1	2/2/2	1/2/1	1/2/1	1/1/1	1.27
5	1/1/1	2/1/2	2/1/1	2/2/2	2/1/1	1/1/1	1.38
6	3/3/3	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1.33
7	2/2/2	1/2/1	1/2/1	1/1/1	1/2/1	1/1/1	1.33
8	2/3/2	2/2/2	1/2/1	1/1/1	2/2/2	1/1/1	1.61
9	3/2/2	3/1/1	1/1/1	1/1/1	1/2/1	1/1/1	1.38
10	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1
Average level of tasks →	1.86	1.666	1.5	1.666	1.4	1	
Overall level = $\dfrac{\text{Average level of tasks}}{6}$ = 1.51							

The overall level of tasks, i. e., 1.51 was by any measures a below par score. The level was quite similar to what was assigned to the tasks designed by *Teacher 1*. Among the average level of tasks, the task on vocabulary (Task 6) stood out as the lowest-rated and it was lower than others by a distinct margin. The tasks for Grammar and Reading with scores 1.4 and 1.5 respectively fared better. But not even a single task came close to the average level.

The average scores assigned to the individual descriptors in the rating scale were also quite low. But the gap among some descriptors was noticeable. Once again, the average score for feedback was the lowest among the scores. In contrast, two other characteristics— one focusing on writing task objectives and the other on the possibility of being integrated into classroom activities— were found to be of average level. The gap of 1.16, between the two extreme scores, was quite large and the gap between the overall average score and the score for the description on feedback was quite significant.

Post-intervention Ability

Teacher 1

The assessment tasks prepared by *Teacher 1* after the intervention were rated and levels were assigned to each task characteristic and in turn, each task. The following table contains the details.

Table 6.5: Post-Intervention Task Levels for *Teacher 1*

Task characteristic	Levels assigned to tasks by 3 different scorers						Average level for each characteristic
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	3/2/2	3/3/3	3/2/3	3/3/3	3/1/2	3/2/3	2.61
2	4/3/4	4/4/4	4/3/3	4/4/4	3/3/3	2/3/3	3.44
3	3/3/3	4/5/5	4/4/4	3/4/4	4/3/4	4/4/4	3.83
4	4/3/4	4/4/4	3/4/4	4/5/4	4/3/3	4/4/4	3.83
5	3/3/3	4/4/4	3/4/4	4/4/4	4/3/4	4/4/4	3.72
6	4/4/4	4/4/4	4/4/4	3/5/4	4/3/4	4/4/4	3.94
7	4/4/4	4/4/4	4/4/4	3/4/4	4/3/3	3/4/4	3.77
8	4/4/4	3/3/3	5/4/4	3/3/3	3/3/3	4/4/4	3.55
9	4/4/4	4/4/4	4/4/4	3/4/3	3/2/3	4/4/4	3.66
10	4/3/3	4/3/3	4/4/4	4/4/4	3/3/3	3/4/3	3.5
Average level of tasks →	3.5	3.8	3.76	3.7	3.13	3.63	
Overall level = $\dfrac{\text{Average level of tasks}}{6}$ = 3.58							

The overall level, indicated by the score 3.58, suggests that *Teacher 1* prepared good assessment tasks for the four skills and two components. Except for a hitch in the task for Grammar, she performed consistently well in all the tasks. Her task for assessing Speaking was assigned the highest score. For individual task characteristics, once again, she performed consistently except for her low scores in writing assessment objectives. There was a clear gap between 2.61, i. e., the average level for writing assessment objectives and 3.69, i. e., the mean average level score for other task characteristics. The teacher got 3.94— her highest score among the task characteristics— in placing the task in real life contexts. This characteristic is one of the most important features of classroom assessment. It shows that the teacher had a good understanding of the concept of 'authenticity'.

The post-intervention tasks prepared by *Teacher 2* were rated and the scores are presented below in tabular form.

Table 6.6: Post-Intervention Task Levels for *Teacher 2*

| Task characteristic | Levels assigned to tasks by 3 different scorers | | | | | | Average level for each characteristic |
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	4/2/4	2/4/3	3/2/3	3/4/4	3/3/3	3/3/3	3.11
2	4/3/3	4/3/4	4/3/4	4/4/4	4/4/4	3/2/3	3.55
3	3/4/3	4/4/4	4/4/4	3/4/3	4/4/4	4/3/3	3.66
4	4/4/4	4/3/3	4/3/3	4/4/4	4/4/4	4/2/4	3.66
5	4/4/4	4/3/4	4/3/3	4/4/4	4/4/4	3/2/3	3.61
6	4/3/4	4/4/4	4/4/4	3/4/4	4/3/4	3/2/3	3.61
7	4/4/4	4/4/4	4/4/4	3/4/3	4/4/4	3/3/3	3.72
8	4/4/4	3/3/3	4/4/4	3/4/3	3/3/3	3/2/3	3.33
9	3/3/3	4/4/4	4/4/4	3/4/4	4/4/4	3/3/3	3.61
10	3/4/4	4/4/4	4/3/3	4/4/4	3/3/3	3/3/3	3.5
Average level of tasks →	3.63	3.66	3.63	3.7	3.66	2.93	
Overall level = $\frac{\text{Average level of tasks}}{6}$ = 3.53							

The overall performance of the teacher was found to be of 'good' level. Among the tasks prepared for assessing the language skills and components, the teacher performed uniformly across tasks except for the one for Vocabulary. The score was significantly lower than all the other tasks. In fact, the difference between the score for Vocabulary task and the mean score for the average levels for the other tasks was 0.72. In case of task characteristics, the teacher's performance in designing assessment tasks was slightly lower in case of writing assessment objectives and maintaining reliability in scoring the task performance. In all other cases, the scores were above 3.5.

Changes in Ability

Teacher 1

The first teacher, with an overall score of 1.55, was assigned a below-average level for her pre-intervention tasks. In contrast, she scored an overall score of 3.58 for her post-intervention tasks on a 5-point scale. So, the overall improvement recorded by the teacher was 40.6%. But this percentage indicates that the teacher, at least at the level of

designing assessment tasks, benefited greatly from the intervention programme and that the intervention programme had some observable impact on the teacher's ability to design classroom assessment tasks. This progress could also be traced for tasks based on individual language skills and components and task characteristics. In her post-intervention tasks, she created descriptive levels for all language skills and components and integrated feedback in those descriptions. There was a great deal of clarity in her presentation of tasks. She also showed adherence to assessment objectives in her tasks.

Teacher 2

The second teacher from the CBSE school also made substantial progress in terms of designing classroom assessment tasks. The gap between his pre-intervention overall score, i.e., 1.51 and his post-intervention overall score, i.e., 3.53 was 40.4%. Before the intervention, he had very little idea about concepts like language proficiency, basic principles of assessment, difference between content and skills, sub-skills of language, importance of classroom assessment, feedback and alternative assessment. It was evident from his post-intervention tasks that he managed to acquire some of the important nuances of classroom assessment and task designing through his participation in the teacher development programme. His growth was consistent across language skills and components and most of the task characteristics.

Overview of Changes in Case 1

There were a lot of common changes traced in the assessment tasks prepared by Case 1 (CBSE) teachers. First of all, the overall increase was around 40% for both of them. Then, both the teachers exhibited some effort in making their assessment tasks learner-centric and tried to utilise some of the alternative assessment methods discussed during the intervention. Thirdly, they seemed to have a little more clarity about the objectives of their respective assessment tasks. Fourthly, they created assessment criteria for almost each task. Though the assessment criteria were far from being perfect, nonetheless, they indicated the teachers' attempt to experiment with the newly learnt ideas. Lastly, their plans about offering feedback to students made more sense.

6.3.2 Case 2

Pre-intervention Ability

Teacher 3

The levels assigned to pre-intervention tasks designed by *Teacher 3* are presented below in tabular form.

Table 6.7: Pre-intervention task levels for *Teacher 3*

| Task characteristic | Levels assigned to tasks by 3 different Scorers | | | | | | Average level for each characteristic |
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	2/1/2	2/2/2	3/2/3	1/3/2	1/3/1	1/2/1	1.88
2	3/2/2	1/1/1	2/1/1	1/3/1	2/3/2	2/2/2	1.77
3	3/2/2	1/1/1	2/1/2	1/2/2	1/2/2	2/2/2	1.72
4	3/2/2	2/1/2	2/1/2	2/2/2	1/2/2	2/1/2	1.83
5	2/2/2	2/1/2	1/1/1	2/2/2	1/2/2	1/1/1	1.55
6	3/2/3	2/1/2	2/1/1	1/2/1	1/2/2	1/1/1	1.61
7	3/1/2	2/1/2	1/1/1	1/1/1	1/1/1	1/1/1	1.27
8	1/1/1	2/1/2	1/1/1	2/2/2	3/2/2	3/1/3	1.72
9	3/1/2	1/1/1	1/1/1	1/2/1	1/2/2	1/1/1	1.33
10	1/1/1	1/3/2	1/1/1	1/2/2	1/1/1	1/1/1	1.27
Average level of tasks →	1.93	1.53	1.36	1.66	1.66	1.43	
Overall level = $\dfrac{\text{Average level of tasks}}{6}$ = 1.59							

With an overall level score of 1.59, *Teacher 3* fell into the category of the two previous teachers who had only fractionally fewer scores. The ability of the teacher to design tasks as displayed through these pre-intervention tasks was less than average for all the four skills and two components. Moreover, the scores for all the task characteristics were also less than 2. However, the range of scores was less wide for all the tasks and task characteristics than the ones found in case of the previous two teachers. To be exact, for the tasks, the range was 0.57 and for the characteristics, 0.61. There was a significant difference between the range scores 0.57 and 1.16 (obtained by *Teacher 2*).

The teacher got the lowest score for the task created for assessing Reading. It was marginally less than what she got for Vocabulary. Her highest score was 1.93 which was in Listening. Among the scores for task characteristics, she got 1.88, her highest score, for writing objectives clearly and appropriately.

154

Teacher 4

The following table contains details about the levels assigned to the tasks designed by *Teacher 4*.

Table 6.8: Pre-intervention task levels for *Teacher 4*

Task characteristic	Levels assigned to tasks by 3 different Scorers						Average level for each characteristic
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	3/2/2	2/2/2	3/3/3	1/2/1	2/3/2	1/2/2	2.11
2	2/1/1	2/1/2	1/1/1	1/2/2	2/2/2	1/2/1	1.5
3	2/1/1	2/1/1	1/1/1	1/2/1	2/1/2	1/1/1	1.27
4	2/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1.05
5	2/1/2	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1.11
6	2/1/2	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1.11
7	2/1/2	1/1/1	1/1/1	1/2/1	1/1/1	1/1/1	1.16
8	2/1/1	1/1/1	1/1/1	1/1/1	2/1/1	1/1/1	1.11
9	2/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1.05
10	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1/1/1	1
Average level of tasks →	1.5	1.2	1.2	1.16	1.33	1.1	
Overall level = $\frac{\text{Average level of tasks}}{6}$ = 1.24							

The overall average level of tasks for *Teacher 4* was 1.24 as highlighted in the above table. This was the lowest score among all pre-intervention tasks constructed by all the six teachers. The low scores were in fact quite consistent for individual tasks designed for each language skill and component. However, though the range was 0.4 for average levels of tasks, it was 1.11 for characteristics of tasks.

The teacher just managed to touch an average level in writing objectives. In all other characteristics, the performance was less than average. Once again, the plan for providing feedback to students was given the lowest rating. It got just 1. The two next best scores, both of them 1.05, were given to two other characteristics- one focusing on giving equal chance to all students to perform and the other, providing information about students' ability to use the targeted language skills.

Post-intervention Ability

Teacher 3

The levels assigned to the post-intervention assessment tasks constructed by *Teacher 3* are presented in the following table.

Table 6.9: Post-Intervention Task Levels for *Teacher 3*

Task characteristic	Levels assigned to tasks by 3 different Scorers						Average level for each characteristic
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	3/3/3	4/3/4	3/3/3	3/3/3	3/3/3	2/3/3	3.05
2	2/3/3	4/4/4	4/4/4	3/3/3	4/3/3	3/4/3	3.38
3	3/2/3	4/3/3	4/3/4	4/4/4	4/4/4	4/4/4	3.61
4	3/2/2	4/3/3	4/3/3	4/4/4	4/4/4	4/3/4	3.44
5	3/2/3	4/4/4	4/3/4	4/4/4	4/4/4	4/3/3	3.61
6	3/3/3	4/2/4	4/3/3	3/4/4	4/4/4	4/3/4	3.5
7	3/3/3	4/4/4	4/4/4	3/4/3	4/3/3	4/4/4	3.61
8	3/3/3	4/3/4	4/4/4	4/4/4	4/4/4	3/4/4	3.72
9	3/3/3	4/3/3	4/4/4	3/4/4	4/4/4	4/4/4	3.66
10	3/3/3	4/4/4	4/3/4	4/5/5	3/4/4	3/3/3	3.66
Average level of tasks →	2.83	3.66	3.66	3.73	3.73	3.53	
Overall level = $\dfrac{\text{Average level of tasks}}{6}$ = 3.52							

The overall level of the assessment tasks was 3.52 for *Teacher 3.* The teacher was found to have constructed 'good' assessment tasks. The only task in which she was assigned a lower level was the one for Listening. The score 2.83 is significantly lower than the overall average, 3.52. In task characteristics, most of the scores are close to 3.52, i. e., the overall level score. However, the first characteristic— focusing on writing assessment objectives— got an average level of 3.05 which was quite lower than the overall average. The previous two teachers also had comparatively lower scores in the same characteristic. The second and the fourth characteristics got 3.38 and 3.44 respectively, and they were much closer to the overall level than the first one.

Teacher 4

The report about the performance of *Teacher 4* in designing assessment tasks after the intervention is presented below in tabular form.

Table 6.10: Post-Intervention Task Levels for *Teacher 4*

Task characteristic	Levels assigned to tasks by 3 different scorers						Average level for each characteristic
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	4/3/4	3/4/3	4/3/3	4/5/4	3/3/3	1/2/2	3.22
2	4/3/4	4/4/4	2/3/3	4/5/4	4/3/3	2/3/3	3.44
3	4/3/3	4/4/4	4/3/4	4/4/4	4/4/4	3/3/3	3.66
4	4/4/4	4/4/4	4/3/4	4/4/4	4/4/4	3/3/3	3.77
5	4/4/4	4/4/4	4/3/3	4/4/4	4/5/4	4/3/3	3.83
6	4/3/4	4/4/4	2/3/3	4/4/4	3/3/3	3/2/3	3.33
7	4/4/4	4/4/4	2/3/3	4/4/4	4/4/4	2/3/2	3.5
8	4/4/4	4/4/4	3/3/3	4/5/5	3/4/4	2/3/3	3.66
9	4/4/4	4/4/4	3/4/4	4/5/5	4/4/4	3/3/3	3.88
10	4/4/4	4/4/4	4/4/4	3/5/4	4/4/4	3/3/3	3.83
Average level of tasks →	3.83	3.93	3.26	4.2	3.73	2.73	
Overall level = $\dfrac{\text{Average level of tasks}}{6}$ = 3.61							

The overall average of post-intervention tasks for the teacher was 3.61, which suggests that the teacher achieved the 'good' level in preparing tasks. He prepared appropriate tasks on most occasions and specifically for Writing, in which he got 4.2— a score indicating excellence. The only problem was with the Vocabulary task in which he got 2.73. This score was too low when compared to the overall level score and the average score for the rest of the 5 tasks. The most important thing is that it fell below the minimum score for the 'average' level.

Among the scores assigned to task characteristics, the first characteristic got the lowest score. Though it did not get a score lower than 3, i. e., the minimum score for 'good' level, the difference between the highest score, i. e., 3.88 and 3.22 was quite substantial. Moreover, the second lowest score was in the characteristic related to authenticity. Since authenticity is considered to be a very important part of classroom assessment, the low scores could be of some concern.

Changes in Ability

Teacher 3

The change in the ability to design effective classroom assessment tasks was significant for *Teacher 3*. Her post-intervention overall score 3.52 was significantly higher than 1.59, the overall score she was assigned for her tasks prepared before the intervention. For all

the skills and task characteristics, there was clear improvement in her average scores. Though she made relatively less progress in designing listening tasks, still she had an improved show. Her progress was also evident from the way she integrated skills and components, made use of alternative assessments, centralized her learners' interests and made her tasks learning-oriented.

Teacher 4

The fourth teacher got the lowest overall score for his pre-intervention tasks. In contrast to that score, which was 1.24, he made a leap of 47.4% to reach an overall score of 3.61 for his post-intervention tasks. He showed improvement in his ability to design tasks for all skills and components and also in adhering to the task characteristics. His post-intervention tasks were integrative, formative and informal in nature. These characteristics were not found in the tasks he prepared before the intervention. In addition, he made use of descriptive grades and rubrics for most of the tasks.

Overview of Changes in Case 2

Not only in terms of overall increase, which was more than 40%, but also in several aspects of CA, the State Board (Case 2) teachers displayed quite a few common changes. When examined closely, these changes indicate that both the teachers brought changes that would help them in carrying out CAs more effectively. Firstly, they adopted an informal approach, moving away from a formal one and tried to assess more than one skill through a single task. Secondly, they wrote assessment objectives a little more appropriately. Then, they drew on developmental criteria for assessing students' performance. Lastly, they thought about focusing on individual learners while offering feedback.

6.3.3 Case 3

Pre-intervention Ability

Teacher 1

The table presented below carries a statistical report about the ability of *Teacher 5* to design effective assessment tasks before undergoing the training programme conducted by the researcher.

Table 6.11: Pre-intervention task levels for *Teacher 5*

| Task characteristic | Levels assigned to tasks by 3 different Scorers | | | | | | Average level for each characteristic |
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	2/1/1	2/1/1	2/1/1	3/2/2	3/3/3	2/3/2	1.94
2	2/1/1	2/2/2	2/1/2	3/2/2	2/2/2	2/1/2	1.83
3	2/1/1	1/2/2	2/1/2	2/2/2	2/2/2	2/1/1	1.66
4	2/1/1	1/1/1	2/1/2	2/1/2	2/1/2	2/1/1	1.44
5	2/1/2	1/1/1	2/1/2	2/1/2	2/1/1	2/1/1	1.44
6	3/1/2	1/1/1	1/1/1	2/1/2	2/1/1	2/1/1	1.38
7	3/1/2	1/1/1	1/1/1	2/1/1	2/1/1	2/1/1	1.33
8	2/1/1	1/1/1	1/1/1	2/1/1	3/2/2	2/1/1	1.38
9	2/1/1	1/1/1	1/1/1	2/1/1	2/1/1	2/1/1	1.22
10	2/1/1	1/2/2	1/1/1	2/1/1	1/1/1	2/2/2	1.38
Average level of tasks →	1.5	1.26	1.3	1.7	1.73	1.53	
Overall level = $\frac{\text{Average level of tasks}}{6}$ = 1.5							

The teacher got an overall average level of 1.5, and her performance was consistently low across the tasks. Her score for the Speaking task was 1.26, the lowest among all and for Grammar, 1.73, the highest average level. The range stood at 0.47. In contrast, the average scores for task characteristics were relatively more scattered. There was almost a gradual decrease in the value of the scores from the first to the tenth characteristic. The best performance of the teacher was in writing task objectives which was rated as 1.94 and the least effective one was providing equal chance to all students to perform, which got 1.22. The range of 0.72 was substantial considering the low scores assigned to the task characteristics.

Teacher 6

The following table provides information about the teacher's pre-intervention performance in terms of designing effective assessment tasks.

Table 6.12: Pre-intervention task levels for *Teacher 6*

Task characteristic	Levels assigned to tasks by 3 different Scorers						Average level for each characteristic
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	2/2/2	2/1/1	2/2/2	2/1/1	1/2/2	1/1/1	1.55
2	3/1/1	2/1/1	3/2/2	2/1/1	2/2/2	2/1/1	1.66
3	3/1/1	3/1/2	3/3/3	2/1/1	2/3/2	2/1/1	1.94
4	3/1/2	2/1/2	2/2/2	2/1/1	1/2/1	2/1/1	1.61
5	2/1/1	3/1/1	2/2/2	2/1/1	1/2/1	1/1/1	1.44
6	2/1/1	2/1/1	2/2/2	1/1/1	1/2/1	1/1/1	1.33
7	2/1/1	3/1/1	2/2/2	2/1/1	1/2/2	1/1/1	1.5
8	2/2/2	2/1/2	1/3/1	1/1/1	2/3/2	2/1/1	1.66
9	2/1/2	2/1/2	2/2/2	2/1/1	2/2/2	2/2/2	1.77
10	1/2/1	1/1/1	1/2/1	2/1/1	1/2/2	3/2/2	1.5
Average level of tasks →	1.63	1.53	2.03	1.26	1.76	1.36	
Overall level = $\frac{\text{Average level of tasks}}{6}$ = 1.59							

Much like other five teachers who participated in the study and designed assessment tasks before and after the training programme, *Teacher 6* did not do anything out of ordinary. The overall average 1.59 gives a fair idea about her ability to construct assessment tasks.

In individual tasks, her scores ranged from 1.26, which she scored in Writing, to 2.03, which she scored in Reading. The range was quite high. Contrarily, the average scores assigned to individual characteristics were a little less scattered. The teacher got 1.94 on the characteristic that was related to the scope of getting the task integrated into classroom activities. The lowest score was 1.33, which was for simulation of real life context in the task.

Post-intervention Ability

Teacher 5

The following table contains the scores assigned to the assessment tasks designed by *Teacher 5* after the intervention.

Table 6.13: Post-Intervention Task Levels for *Teacher 5*

| Task characteristic | Levels assigned to tasks by 3 different Scorers | | | | | | Average level for each characteristic |
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	3/3/3	3/4/4	4/4/4	4/4/4	4/3/3	3/2/3	3.44
2	4/4/4	4/4/4	4/4/4	4/4/4	4/4/4	4/3/3	3.88
3	4/4/4	4/4/4	4/4/4	4/4/4	5/4/5	4/3/4	4.05
4	4/4/4	5/4/5	4/4/4	4/4/4	4/3/4	4/3/4	4
5	4/4/4	4/4/4	4/4/4	4/4/4	3/3/3	4/3/4	3.77
6	4/4/4	5/4/5	4/4/4	4/4/4	4/3/4	4/3/3	3.94
7	4/4/4	4/4/4	4/4/4	4/4/4	4/3/3	4/3/4	3.83
8	4/4/4	4/4/4	4/4/4	4/4/4	4/2/3	4/3/4	3.77
9	4/4/4	4/4/4	4/4/4	4/4/4	4/2/3	4/3/3	3.72
10	4/5/4	4/4/4	4/3/4	4/5/5	5/4/4	4/3/3	4.05
Average level of tasks →	3.93	4.1	3.96	4.06	3.6	3.43	
Overall level = $\dfrac{\text{Average level of tasks}}{6}$ = 3.84							

The overall average score of 3.84 suggests that the teacher did design effective assessment tasks and had a good understanding of the principles of assessment. Her scores for individual tasks were quite close to 4 for the four language skills, the best one being 4.1 for Speaking. For Grammar and Vocabulary, however, the scores were lower. In fact, the lowest score was assigned to the task for Vocabulary. Once again, the lower scores in Vocabulary continued to appear.

The teacher's performance in task characteristics was somehow consistent except the visibly low score in writing objectives. Among the rest of the characteristics, the lowest score was 3.72 which was not substantially lower than 4.05. The scores also reflected the teacher's understanding of the basic principles of assessment and other aspects of assessment like feedback.

Teacher 6

The tasks designed by *Teacher 6* were rated. The details about the scores assigned to the tasks are presented in the following table.

Table 6.14: Post-Intervention Task Levels for *Teacher 6*

Task characteristic	Levels assigned to tasks by 3 different Scorers						Average level for each characteristic
	Task 1	Task 2	Task 3	Task 4	Task 5	Task 6	
1	4/3/3	4/3/4	4/4/4	4/4/4	4/4/4	4/4/4	3.83
2	4/2/3	4/4/4	4/4/4	4/4/4	5/4/5	4/4/4	3.94
3	4/4/4	4/4/4	4/3/4	4/5/4	4/4/4	4/4/4	4
4	4/2/3	4/4/4	4/4/4	4/5/5	4/5/4	4/4/4	4
5	4/2/3	4/4/4	4/4/4	4/5/5	5/5/4	4/3/4	4
6	4/2/3	4/4/4	4/3/4	4/5/5	5/5/5	4/3/4	4
7	4/3/3	4/4/4	4/4/4	5/5/5	5/5/5	4/3/3	4.11
8	4/2/3	4/4/4	4/4/4	5/5/5	4/5/5	4/4/4	4.11
9	4/2/3	4/4/4	4/4/4	5/5/5	5/5/5	4/4/4	4.16
10	4/2/4	4/4/4	4/4/4	5/5/5	4/5/5	3/4/3	4.05
Average level of tasks →	3.2	3.96	3.93	4.63	4.6	3.8	

Overall level = $\dfrac{\text{Average level of tasks}}{6}$ = 4.02

The teacher just touched the overall level of 'excellent' in designing assessment tasks. Her overall score, however, does not really show her performance in all the individual tasks. The gap between the highest and the lowest score was 1.43 which was a considerable one by any standard. Moreover, the scores for other tasks were also quite scattered. In contrast, the scores for task characteristics were not as varied and scattered. All the scores were very close to the overall score of 4.02. However, it was interesting to see that the teacher got her lowest score in writing assessment objectives.

Changes in Ability

Teacher 5

The changes exhibited by *Teacher 5* in terms of ability to design appropriate classroom assessment tasks after the intervention were pleasantly surprising. She recorded an amazing 46.8% growth over her overall score for her tasks by moving from 1.5— her overall pre-intervention score, to 3.84— her post-intervention score. Her improvement was evenly spread across all the skills and components and all the task characteristics. Considering that she was more than 60 years of age and was not required to learn much about classroom assessment by her school management, she improved brilliantly. The most distinctive change was found in the assessment criteria/rubrics she prepared for

each task. The descriptors were to a great extent appropriate. The tasks bore the direct impact of the intervention content.

Teacher 6

The progress made by *Teacher 6* stood out among all the participant-teachers. She moved from 1.59, her pre-intervention overall score to 4.02, her post-intervention overall score. The percentage of her growth was 48.6%, the highest among all the teachers who were part of the intervention programme. Like *Teacher 5,* her colleague, she showed great change in her approach to classroom assessment after the intervention. In her post-intervention tasks, she created proper level descriptors, wrote clear assessment objectives, tried to provide feedback to students through the rubrics, integrated the assessments with classroom learning and made use of situations in the classroom to collect information about students' progress.

Overview of Changes in Case 3

Both the ICSE (Case 3) teachers recorded an increase of around 47% in their scores assigned to the assessment tasks designed by them. Moreover, their post-intervention tasks were a little more learner friendly than the pre-intervention ones. However, the tone of the tasks remained much the same. The teachers adopted a descriptive approach to writing objectives of the tasks. In addition, the assessment criteria designed for each task used somewhat appropriate descriptors. Even their feedback plan incorporated peer feedback along with that by the teacher. Much like the change in beliefs, these changes did not reflect actual change in the teachers' approaches to teaching and assessment. Once again, the institutional and board assessment policies might have influenced the attitude of the teachers.

6.3.4 Cross-case Analysis of Changes

A comparative analysis of changes found in the teachers' ability to design assessment tasks, write assessment objectives clearly and accurately, create and use assessment criteria and develop a feedback plan is presented in tabular form below. The table provides a good idea about how the three cases fared against each other.

Though there is not much difference among the cases in terms of the overall progress made, subtle differences were recorded in some of the main aspects of CA. In assessment task design, the three cases changed in different directions. Whereas Case 1

teachers tried to move from teacher-centred to learner-centred tasks and made use of alternative assessment methods, Case 2 teachers changed their approach of assessing one skill at a time to an integrated-skills approach and created informal assessments, and Case 3 teachers moved from very little learner-involvement to some learner-involvement and continued   to   have   a   formal   tone   in   their   tasks.

Table 6.15: Cross-case Analysis of Changes

Case	Overall Increase (in percent)	Assessment Tasks	Writing Assessment Objectives	Use of Assessment Criteria	Feedback Plan
1	40	• more variety in tasks moving towards learner-centredness • alternative assessment methods used	• better awareness of sub-skills of language • used exact words as used during the TD tasks	• from very little to some rudimentary awareness about how to develop assessment criteria	• more focused • utilized newly learnt ideas creatively
2	43	• single skill to integrated-skills approach • formal to informal	• from broad to narrow and specific	• from very little to some awareness of assessment criteria • process-oriented and developmental criteria	• from overemphasis on accuracy to emphasis on relevant areas • from complete group feedback to individual feedback
3	47	• from teacher-dictated to learner-oriented • formal tone does not change much	• better awareness sub-skills, clearer and descriptive objectives	• much better understanding of assessment criteria • appropriate description of levels • learner-centred approach to criteria development	• focus shifted from mere 'corrections' to plans to improve learning • from teacher feedback to 'teacher + peer' feedback

In all three cases, teachers became more aware of sub-skills of language and wrote assessment objectives more effectively and the direction of change was quite similar. The assessment objectives written by Case 3 teachers were a little more convincing than those by the teachers belonging to the other two cases.

As evident from the table (Table 6.15), the cases explored differently when it came to creating assessment criteria. While Case 1 teachers exhibited some rudimentary awareness of assessment criteria, Case 2 teachers made an attempt at developing a set of process-oriented and developmental assessment criteria for several tasks. In contrast, Case 3 teachers refined their descriptors and tried to make their level descriptors learner-friendly.

There were some differences observed in the way the teachers in all three cases planned to offer feedback to their students. Case 1 teachers tried to use the newly learnt ideas about feedback creatively and became more focused. Case 2 teachers moved away from overemphasis on accuracy to sub-skills being assessed and total group feedback to individual feedback. Case 3 teachers also moved from their overriding concern about accuracy before intervention and planned to offer directions to their learners so that they could learn to use the focused skills more effectively. In addition, they also thought of using peer feedback as an option.

Data were collected to corroborate the findings obtained through the evaluation of simulated CA tasks. The data, which are presented in the following section, comprised classroom observation reports and information acquired through interviews with the teachers.

6.4 Changes in Teachers' Assessment Practices: Interviews and Classroom Observations Data

Since the main thrust of the research was to explore the relationship between teachers' CLAL and their assessment practices, evidence regarding teachers' assessment practices was collected through different means. The teachers were asked to design assessment tasks before and after the intervention in a simulated setting. There were pre- and post-intervention classroom observations and interviews for each individual teacher. This section presents an analysis of the interview and classroom observation data for each teacher. An overview of each case along with a cross-case analysis also forms parts of this section.

6.4.1 Case 1

Teacher 1

Interview data

The first teacher did not want to spend much time on thinking about assessment and designing assessment tasks. She was burdened by workload at the school and was not willing to explore new methods of assessment. She also informed that the school administration had an assessment policy which was rigidly driven by a 'misinterpreted' version of CCE. She wanted the school administration to be aware of classroom assessment and its benefits.

She followed the teacher support books provided by CBSE for conducting internal assessments. However, she thought the book did not provide enough examples and guidance regarding how to employ different alternative methods of assessment. So she had to depend on old question papers or textbooks for constructing assessment tasks for her students. Though she gave oral and corrective written feedback to students on their performance in the internal assessments, she thought she did not know how to provide individual feedback to students regularly and effectively. She also stated that she did not know how to evaluate and improve the quality of assessment tasks. She strongly felt that English teachers in schools should be given regular opportunities to meet experts in language assessment and discuss their problems related to classroom assessment with them. It was difficult for her on many occasions to find out whether the way she was assessing her students was appropriate or not.

Classroom observation before intervention

On the first day of observation, the teacher was conducting a formative assessment. She was assessing the students' listening and speaking skills. Students were told about the assessment well in advance. However, she did not tell students about the assessment criteria on the basis of which their performance in listening and speaking assessments would be graded. That seemed a little unfair and indicated that she did not know much about assessment criteria and their utility. She spent two periods, one of which was observed, for carrying out the assessments. For the listening part, she used an audio clip, in which there was a speech by an Indian, and a blank-filling activity. She asked the students to fill in the blanks individually while listening to the audio clip. She played the

clip twice. She conducted the speaking assessment after completing listening. Students were given three topics- "Pollution", "Children's Day" and "Diwali". They were asked to prepare and speak on one of the topics. The teacher along with another English teacher from the same school graded them. The speaking assessment could not be completed in one period. It was continued in the next period too. However, the teacher did not have any plans to provide individual feedback to students. After the examination, she asked them to work hard on the listening and speaking skills and improve their performance in the next formative assessment. While the listening assessment task was acceptable, the speaking task did not seem to serve the intended purpose. Most of the students might have memorised what they wanted to speak and reproduced the same in the class. The teacher did not possibly have enough knowledge about assessment of speaking skills to think about a more valid way.

There was no formal assessment happening on the second day of observation. The teacher taught Reported Speech to class IX students. She wrote the grammatical rules of reported speech and an example for each rule on the blackboard. Then she asked the students to convert a set of sentences in direct speech to reported speech. She had a classroom discussion in which she discussed the answers and all the students were asked to correct their answers if they had written wrong ones. She was sure that those who had problems with using reported speech could learn if they followed the rules and did practice exercises at home. Before leaving the classroom, all the students were asked to submit their answer books to her. It was quite obvious that the teacher wanted her students to learn how to use reported speech in communicative situations. However, she did not provide adequate context for them to learn the targeted item. It could be due to her either negligence or ignorance.

On the third day of observation, the teacher taught a lesson from the Main Course Book. First, she read the lesson aloud to the students and after reading a paragraph, she asked a few comprehension questions to the class. Some of the students responded with correct answers. She repeated each correct answer and moved ahead with the lesson. She did not try to ask questions and elicit answers from quite a few students. Moreover, when students could not answer any of the questions asked by her, she directly gave them answers. After the class, she informed the researcher that she never forced or 'tortured' any of the 'weak' students by asking them questions in the class and that she meticulously corrected their mistakes in the answer books and gave them correct answers quite

regularly. Once again, much like what she did in her previously observed lesson, she had lines drawn between 'right' and 'wrong' answers. Although it was a good move to repeat the correct answers, not involving students who could not give 'right' answers was surely not part of good practice. By not involving them, she lost an opportunity of teaching and promoting learning. Her inability to utilise classroom questioning for learning purposes was apparent. Moreover, she did not offer any constructive feedback to students.

Classroom observation after intervention

During all three days of classroom observation, the teacher did not have any prescribed or formal test for her students. She taught Reading, Vocabulary, Grammar and Listening in these three classes. In the first class, she taught a text from the main course book. She made use of a variety of questions that included 'Yes-No', 'Wh-' and 'True-False' types. To ensure that students had read the text, she mixed inferential, global and extrapolative questions along with factual questions to students. Most of the factual questions were targeted at slow learners. On the other hand, the other types of questions were asked to average and fast learners. There was some amount of improvement in terms of the way she used questioning for teaching. It is quite possible that she made an effort to try out asking a variety of questions and involving all the learners after undergoing the TD sessions in which there were explicit discussions on the same. However, she may or may not have realised the benefits of it in the CA context. In the second class, she started with a listening comprehension task. The students were asked to fill out information on a flowchart while listening to a conversation. Two students were asked to play roles and enact a situation of conversation that happened between a stranger in a city and an auto driver to whom he was asking directions. Here more than anything, the communicative context used for the task indicated some development. After the listening activity, she taught the use of modal auxiliaries by using role playing activities. Each pair was evaluated by their peers on the basis of a pre-designed rubric. The teacher monitored the activities and gave her feedback at the end of the class using her notes which she took during the activities. In this activity, once again, the context was adequate. In addition, the teacher employed rubrics and notes to offer feedback to her students.

In the last class, she brought a newspaper article which was similar to the prescribed text in terms of genre and difficulty level. Students were divided into pairs and all the pairs were asked to read the text silently, underline the words and phrases they found difficult and guess a possible meaning of the words and phrases in the given text. The teacher

elicited the difficult words, wrote them on the black board, asked each pair to tell the meaning they guessed and initiated a discussion on each word and phrase. In the process, she taught them how to guess meanings of words, evaluate peer responses and make use of guessing while reading. While the silent reading implied a better understanding of language skills and the concept of validity, the teaching of guessing as a skill showed the teacher's progress in classroom questioning and attempt to make learners independent.

Changes in assessment practices: observed and self-reported data

Although not much of direct information about the teacher's assessment practices was collected through observation of classes, still the observed changes in the classroom are relevant to the main focus of the study. Before the intervention, *Teacher 1* did not use any written assessment criteria for her classroom assessment. She did not involve students in the process of assessment and her teaching was not much student-centric. She also seemed to be unaware of how to design assessment tasks for speaking, and did not use suitable communicative contexts to teach and provide constructive feedback to individual students. While teaching grammar, she asked her students to complete an exercise by applying the rules she mentioned earlier on their own. She gave away the right answers leaving little challenge and scope of learning for students. It indicated that she did not have much knowledge about how assessment can be used for promoting learning of language skills in the classroom. In the last class, she assessed her students' reading comprehension skills after practicing loud reading. Further, she asked very limited type of questions which could be a hindrance to obtaining adequate information about students' language ability.

Quite a few things were found changed during the post-intervention observations. She employed a variety of questions—in terms of form and content— and tasks (e. g. flowchart and conversation), integrated skills, tried some innovation by using role-play techniques to teach modal auxiliary and newspaper article to teach reading comprehension, made her classes engaging and learner-centric, drew on rubrics, taught language skills in communicative contexts, encouraged peer-evaluation and made use of classroom observation and notes to provide individual feedback to students. She also exhibited a much better understanding of skills when she asked students to read silently and guess meanings of words in the given contexts. The above-mentioned changes suggest that the intervention programme could have had some impact on the teacher's

approach to teaching, which in turn, might lead to better classroom assessment as it was supposed to be integrated with teaching and learning in the classroom.

The teacher talked about her assessment practices and also showed some of her question papers during the pre- and post-intervention interviews. She informed the researcher that she had learned to create better assessment tasks during the intervention, became aware of how to apply the basic principles of assessment while designing tasks, learned about formative methods of assessment like portfolio and observation, offer effective feedback to individual students without spending much time on it, and evaluate and improve already-used assessment tasks. Though the school wanted her to design certain type of question papers for the formative tests, she felt that she would be able to convince the management about using portfolios, observations and peer-evaluation for classroom assessment purposes. She showed some question papers which had been designed by her with her colleagues before the intervention and claimed that she could change the paper into a much more effective one because of her newly gained knowledge during the intervention programme. She was more independent and confident about her assessment tasks. She also reported that she was going to assess her students informally throughout the academic year during normal classroom practices, and assign them a final grade.

Teacher 2

Interview data

The second teacher was too worried about teaching methodology and his own English language skills to think about assessment. He had very little idea about methodology for teaching language skills. He never tried to design question papers for internal assessment. Whenever he was asked to assess, he asked his students to develop a project on one of the well-known public figures in the world. He never stressed the use of language as the main assessment criterion for those projects. In fact, he did not know about the utility of project work for language learners.

Apart from having very little idea about assessment of language skills, the teacher never gave feedback to his students about their performance in assessments and did not try to evaluate the effectiveness of assessments. He had little idea about how to interpret the examination scores. He strongly believed that he needed more professional help to improve as a teacher and also that the school administration should arrange for the professional development of teachers.

In the beginning, the second teacher was not willing to let the researcher observe his classes. Later, he reluctantly agreed. Also, he never shared the answer scripts of his students with the researcher. In the first class observed by the researcher, he taught a lesson from the Main Course Book in class VIII. He asked each student to read aloud a paragraph from a piece of prose. He explained the meaning of each paragraph before moving on to the next student. Sometimes, he asked a few questions on the text to his students and also enquired with them whether they had understood the text well. In addition, he wanted his students to pronounce each word correctly while reading. After that class, he informed the researcher that he wanted his students to have good pronunciation so that they could read well. It was quite obvious from the teacher's use of loud reading and emphasis on pronunciation for improving his students' reading ability that the teacher had very little understanding of language skills and how they are taught.

In the next class, the teacher conducted a class test. It was formal and all the students were notified about the test. They were given a set of questions on several items of grammar including Articles, Passive Voice and Subject-Verb Agreement. There were blank-filling, error-correction and sentence-conversion questions used by the teacher for that paper-pencil test. It covered the grammatical components already taught by the teacher in the classroom. He asked them to complete the test in twenty minutes. Then he gave the correct answers to the class and the students made necessary corrections and scored their own answer scripts. Before leaving the class, he praised those who got high scores in the test. On the one hand, the teacher did what was commonly practised by his colleagues in the school for 'formative tests', on the other, his test did not seem to serve the purpose, of which he displayed little understanding. The discrete-point test of grammatical items employed by the teacher indicated the teacher's lack of knowledge and skills required to carry out CAs.

In the third class, the teacher prepared his students to take up a small project work in which they had to collect information about a famous personality and develop a project on it. The teacher instructed them to work individually, be original in their presentation, limit the number of pages to maximum 10 pages, make it look beautiful with pictures and colours and submit the project within a month's time. He informed the researcher that he would consider the project under formative assessments and grade students. He never told the students clearly about the criteria of assessment to be followed for the evaluation

and grading of the project work. When used appropriately, project-based assessment can be an effective tool in CA/FA. However, the teacher, though made a good attempt by employing it, did not seem to be aware of how to utilise it. From his emphasis on "making things beautiful", it could be deduced that he was not sure about the objective/s of the project.

Classroom observation after intervention

After the intervention, the teacher did not have the reluctance which was obvious during the pre-intervention observation. In his first class, he taught a poem from the main course book. He wrote a global question on the black board and asked the students to answer the question after listening to the recitation of the poem by him. The teacher's approach here was systematic and purposeful. Then he divided the class into small groups, asked each group to write a few questions on a particular stanza and then respond to questions posed by other groups. At the end of the class, he asked his students to write the answers to questions given at the end of the lesson. The teacher seemed to have a plan here to facilitate learning. Moreover, his understanding of language skills in this class was better than what was observed in his pre-intervention classes.

In his second class, he discussed the answers to questions given at the end of the poem-lesson with his students. It was the continuation of the class observed earlier by the researcher. He allowed almost every student to respond to his question and offered feedback orally after allowing their peers to correct them wherever necessary. Finally, he asked them to write the summary of the poem as homework. One important thing noticed in the class was the teacher's experiment with peer-correction.

In his last class, he completed a set of reading comprehension and vocabulary exercises mentioned at the end of a prose lesson. He read out one question from the book and asked individual students to respond to questions. Whenever somebody failed to respond correctly he asked the same question to others. At times, he gave the right answers to students when he could not elicit them. Once again, the teacher was trying to involve individual learners. Though the teacher did not employ many leading questions to elicit answers, he gave every learner opportunity to respond to the questions.

There were some important changes observed in the teacher's approach to teaching after he participated in the teacher development programme. Before the intervention, the teacher was seriously unaware of teaching objectives and his ignorance was evident in the way he practiced loud reading and emphasized pronunciation to improve reading comprehension skills of her students. He used a discrete-point test and attached too much importance to scores. Both these things are often discouraged in classroom assessment. Though he assigned some project work to his students as a part of assessment, his inability to differentiate between content and skills took the entire process off the track and language learning almost vanished from the picture. Some of the above-mentioned points like awareness of teaching objectives, using communicative contexts, understanding of formative assessment, language proficiency, knowledge about language skills and sub-skills, etc. are crucial to effective assessment practices. During the post-intervention observation of classes, the teacher displayed slightly better knowledge of teaching objectives which was evident from the way he exploited a poem for teaching listening comprehension. He also tried to involve his students by posing different types of questions according to their level. Moreover, he encouraged peer-correction and feedback in the class. He integrated reading and writing skills by asking students to write a summary of the poem. Also, by paying attention to individual students and giving them opportunities to use language skills, he showed a much better understanding of language proficiency. Though these changes were not really many, still they could be interpreted as part of the teacher's developing process in understanding the components of CA discussed during the TD programme.

The teacher talked about his assessment practices during the interviews. During the pre-intervention interview, the teacher reported being worried about his lack of knowledge in teaching methodology. It also became clear that he never conducted an internal assessment because he did not know how formative assessment was different from a term-end examination. What he did for internal evaluation was a project, which was a fun-filled activity for the students. But he did not know how to assess language skills through projects. In addition, he did not offer feedback to students because he thought it would consume all his classroom time. However, he showed a great deal of motivation to develop his ability in assessment. After the intervention, when he was interviewed he said that he was feeling "more powerful and competent" after undergoing training in

assessment. He stated that he would try to design internal assessments with his colleague and create rubrics for assessing language skills through project work. He also felt that he could offer feedback to students without spending much time on it. Whether he put all that he talked about into practice could not be captured due to constraints of time, but the teacher's firm assertion could be interpreted as something positive. This could be an initiation into better assessment practice.

6.4.1.1 Influence of Pedagogic Context and Personal Factors on Changes

The CBSE assessment policy, school assessment culture and the teachers' assessment abilities and beliefs must have influenced the changes in the teachers. As per the CBSE assessment policy, teachers need to integrate assessment with teaching. They need to find out how students are making progress and about students' problems through formative tests. Use of a variety of assessment methods is also suggested. Both the teachers were found making efforts to use portfolio evaluation, observation and peer-evaluation in their respective classrooms. Also, they created rubrics for providing feedback to their students. These things are prescribed and mentioned in the assessment manual of CBSE. However, they were thinking about discussing with the school management the necessity to change the existing traditional test system for formative tests. Another instance of the impact of the school management was evident in the way they wanted to jointly work to design formative assessment tasks. The management must have allowed them to do that. In addition, the busy schedule at the schools made them think about finding time-saving ways to offer feedback. Along with these, the personal abilities and beliefs also had strong influence on the changes. The first teacher was a little more enthusiastic and self-aware and had some idea about CA. These factors also helped her to utilize the new knowledge and skills. Informal assessment, developmental grading, proper use of a newspaper article, teaching of guessing as a skill, etc. reflected her personal ability and beliefs. Her score for the PA section in the beliefs questionnaire also indicated the change in her attitude. In contrast, the second teacher, who wanted to know more about methods of language teaching, took more interest in changing his approach to teaching. His use of a poem to teach listening skills, use of a variety of questions and providing opportunity to his students to use English in the classroom, etc. were indicative of his newly developed awareness about skills and sub-skills. The board or the school had very little influence on these changes. He was also a little more worried about providing feedback to students and thus, wanted to try out using rubrics for the same purpose.

6.4.2 Case 2

Teacher 3

Interview data

The third teacher reported that she was unhappy with the government policy of introducing a new assessment policy without offering any training to teachers. She said that she had to follow the new assessment 'rules' because the school administration wanted her to. She did not know how the new assessment policy would lead to better learning. When asked about how she prepares her internal question papers, she informed that she took help from question banks and other teachers in the school while preparing question papers. She did not know much about alternative methods of assessment, how formative tests were different from examinations they used to have, how to interpret students' scores and evaluate assessment tasks. She found providing individual feedback too time-consuming and was "seriously unhappy" with her students because they did not bother to make use of her corrective feedback on writing.

Classroom observation before intervention

The third teacher was very comfortable with observation of her classes. She conducted a formative assessment during the first class. Her students knew about that "slip test" in which she included both vocabulary and items of grammar from some of the recently taught lessons. In the vocabulary part, she gave a list of ten words and asked the students to write the meaning of those words and use them in sentences. In grammar, she gave her students questions on Reported Speech. There were ten sentences in Direct Speech and they were instructed to convert those independent sentences into Reported Speech. At the end of that period, she collected all the answer scripts from the students. She informed the researcher that she would score all the scripts and return them to the students and also that she would discuss the answers to the questions in the class on the same day. The way the test was conducted indicated that the teacher did not see any difference between summative and formative tests. There was little evidence to claim that she wanted to obtain concrete information about students' progress and problems through the slip test.

In the second class, the teacher tried to help her students with the exercises given at the end of a lesson from the text book. She read aloud each question and asked the students

to respond to it orally. Whenever she got right answers, she repeated the answers and asked the whole class to write down the answers. But when she did not get right answers, she dictated the right answers on her own. She continued in a similar manner and completed the exercises on reading comprehension, vocabulary and grammar. However, the portions on writing were given as homework. It is possible that the teacher did not know much about making use of a variety of questions to elicit answers and involve students in the process. Again, her emphasis on 'right' answers was not a learner- or learning-friendly approach.

On the third day, the teacher came to the class with students' answer scripts— used for the slip test. She had already scored those scripts. After returning the answer scripts to students, she discussed the questions asked in the slip test and tried to elicit correct answers from students. When they could not provide correct answers, she gave them the correct answers. She showed her unhappiness with half of the class because they had fared poorly in the test. She asked them to work hard and practice with more exercises. When the researcher enquired with her about the reason for poor performance of students, she explained by saying that those students belonged to very poor backgrounds, got little encouragement at home to study and did not pay much attention to studies even when they were in the school. She might be right about the background of the students being a reason for their poor performance. However, she did not give any hint about plans to reconsider her approach to teaching and improve students' learning.

Classroom observation after intervention

In her first class that was observed, the third teacher was teaching her students how to describe places. She created a mind map by asking leading questions to students and eliciting answers from them. She added a necessary set of vocabulary to some of the points mentioned in the mind map. Then she made use of a picture and encouraged students to describe it using the vocabulary set. The students responded in groups and completed the description of the picture which contained a place. Then she asked her students to work in groups of three and describe a place/lane/area/locality in writing. Some groups managed to write the description, some could not. She helped the struggling groups by offering them a few sentences. At the end of the class, all the students were asked to describe a place in writing and submit it to her in the next class. In this class, the teacher utilized questioning as a tool to guide students and group-work

to promote learning. Moreover, she managed to confine her focus to 'describing places'- which was what she was trying to teach.

In her second class, she started by collecting answer scripts containing the description of a place. She informed the researcher that she wanted to see how peer correction works. Then she made students exchange their answer scripts with each-other and comment on their peers' writing. She wrote three criteria— completeness of the description, organization and readability, and asked students to keep the set of criteria in mind while responding to their peers' scripts. She collected the answer scripts from the students at the end of the class and looked at some of the scripts while coming out of the class. She was not totally happy with peer correction as an option and told the same to the researcher. However, her attempt to try it in her classroom was a positive sign. Also, she made use of assessment criteria. Though she could have described the criteria better, considering it was an initiative, it can be considered as a positive change.

In the last class observed by the researcher, the teacher started the class by distributing the answer scripts from the previous class. She gave some oral feedback to each student in addition to the written corrective feedback she provided them on their answer scripts. She asked them to look at her comments and the comments made by their peers and resubmit a draft after making necessary corrections. She took one of the scripts, read it out to the class, also read the comments and told her students how they could improve the draft and write a better description. Then she read out one more along with comments and asked students where and how improvements can be made in the script. The teacher did a fairly good job in this class by providing students with constructive feedback and making them comment on each other's writing. Moreover, multi-drafting, a process approach to teaching, was a good move. She went beyond 'wrong-right' answers and offered students scope to improve their writing.

Changes in assessment practices: observed and self-reported data

The third teacher was found to have a relatively better understanding of things related to teaching and assessment after participating in the TD programme on assessment. Earlier, the discrete-point test of grammar and vocabulary given to students was an indicator of the fact that she did not exploit the opportunity to obtain information about her students' learning through the 'formative test'. Next, she made all her students write a single answer to each reading comprehension question. She was not ready to either

accept a variety of answers or spend time on going through multiple responses or both. The way she spoke about the poor performance of her students suggested that she had very little idea about how to interpret and use the results of assessment. She could not think beyond 'good' and 'bad' for her students' test scores.

In the post-intervention observation, however, she displayed an improved understanding of classroom assessment. She adopted a systematic approach to teaching her students how to describe places and provided adequate support to slow learners. A similar sensibility and approach to assessment can be quite rewarding in the sense that all kinds of students can have scope to perform. Another encouraging sign was that she put peer-correction and evaluation criteria to immediate use though she was not fully convinced that peer correction would lead to better learning and performance. A few more perceptible changes included peer correction, evaluation criteria, use of written and oral feedback to individual students and positive conviction about evaluation as a process.

The teacher realised a few changes in her assessment practices and reported the same during the post-intervention interview. Contrary to her pre-intervention mistrust in formative assessment, she thought she could use "formative tests" for keeping track of students' progress. She believed that her awareness of alternative methods of assessment, sub-skills of major language skills, variety of tasks and question-types, etc. would allow her to become a better assessor in the near future. She also informed that she had started using rubrics to provide individual feedback to her students. But she said it would take time for her, her students, the school management and parents to understand language proficiency and test scores well.

Teacher 4

Interview data

The fourth teacher was not quite aware of how the new assessment policy would help in supporting learning in the classroom. He was of the opinion that formative tests unnecessarily burden students and teachers with more examinations. Though students might work hard throughout the year, he thought, such a policy does not make learning burden-free. He prepared his internal assessment question papers and arranged a project work for his students along with his colleagues. He never tried anything other than paper-pencil tests and had no idea how a project work could be properly done to promote language learning.

Apart from regular check of students' answer books, he never offered any individual feedback to students. He informed that he did not know how peer feedback can be used for promoting learning. So he never tried that. He was happy with the previous question papers he prepared and did not want to change things much because he was afraid of the school administration and not sure about his knowledge about the appropriateness of tasks.

Classroom observation before intervention

The researcher knew the fourth teacher very well even before taking up the study. The teacher was quite relaxed about the observation of his classes. In the first class, he focused on a story from the class VIII textbook. At first, he read aloud the first paragraph from the story and then asked a student to read the same passage loudly. Then he asked a few very short comprehension questions to his students. When they had trouble in answering the questions, he answered them. He covered the whole story in a similar manner. Also, he used Telugu to explain meanings of words when students had difficulty in understanding any word. Finally, for homework, he asked his students to write answers to the comprehension questions given at the end of the lesson. He justified his belief in loud reading to the researcher. He believed that loud reading would help his students improve their pronunciation, comprehension and reading ability. One could see how his lack of knowledge about language skills and teaching methodology affected his teaching.

In the second class, the teacher conducted a "unit test". It was a one-hour examination comprising reading comprehension questions based on a text from the prescribed textbook, one informal letter, a poster and three questions each on vocabulary and grammar from the recently completed lessons in the textbook. It was a plain paper-pencil test. Students, who could not complete their writing in the given time, were given a little extra time. The teacher informed the researcher that he would score the answer scripts and keep a record of the marks scored by the students in the test and give correct answers to students in the next class through a classroom discussion of the question paper. Except for the extra time allowed to students to complete writing the test, most of the things did not serve the teacher's intended purpose. Further, there was plenty of evidence to doubt the teacher's understanding of the objective/s of the 'unit test'.

The teacher discussed the question paper used for the unit test in the third class observed by the researcher. He returned the answer scripts of students before discussing the questions and tried to find out how the students had performed in the test. First he asked the students who had scored more than 20 marks to stand up. Then, he lowered the score to 15 and subsequently to 10. He asked everyone to work hard so that they could do better in the final examination and had a few words of advice in Telugu for poor performers in the examination. He went on to discuss all the questions in the class. He told the researcher that it was important to keep a record of the students' scores because the school needed to maintain an account of the scores and he would know if students were improving their performance. This was an indication of what he made out of the scores. He failed to see the problems students were facing in the areas covered in the test.

Classroom observation after intervention

The fourth teacher completed a prose lesson which he started in his last class. While helping them to understand the text, he drew the attention of his students towards grammatical points like prepositions, tense forms and punctuation in the text. He asked his students questions on these points and in the process, tried to make them understand the use of these items in context. After completing the text, he asked them to do a few exercises which included reading comprehension tasks and tasks based on grammar and vocabulary. The questioning strategy and the efforts to exploit the text for developing more than one skill were two positive changes observed in this class. Considering that the integration of skills was discussed during the TD programme, this change meant a lot.

In the second class, *Teacher 4* prepared his students for a project work which required the students to collect information about a famous personality and write systematically about them. The teacher emphasised collecting appropriate information, writing creatively and providing personal comments on the life of the personality. He also told his students about the assessment criteria against which their projects would be assessed. A few students had a few doubts about the assignment and the teacher responded to all the queries. Though project-based assessments were not exclusively discussed during the TD programme, the teacher managed to improve the regular practice of project-based assessments by setting concrete objectives and incorporating a set of assessment criteria in it.

In his third class, the teacher taught how to transfer information from a reading passage to a flow chart. The activity was based on a text which had already been taught by the teacher to the class. The students were divided into groups and asked to complete the flow chart using the information from the text. The teacher monitored the class and helped the groups using Telugu, i.e., the mother tongue of the students, whenever it was necessary. At the end of the class, the teacher initiated a discussion in which he encouraged students to participate while cross-checking the accuracy of the information used to complete the flow chart. The employment of the flow chart indicated the teachers' improved understanding of skills and task construction; and his use of students' mother tongue was judicious. Even his effort to utilize group discussion as a feedback strategy was a constructive move.

Changes in assessment practices: observed and self-reported data

Before participating in the LAL development programme, *Teacher 4* had very little understanding of sub-skills of language. He had problems in setting teaching objectives and translating those objectives into lessons. The learners did not get much scope to perform in the assessment conducted by him. This could be further related to a weak understanding of the concept of validity in language assessment. He could not distinguish between assessing content and skills and perhaps, that is the reason why he used a passage from the prescribed textbook to assess his students' reading comprehension skills. It was almost obvious that he did not know how to interpret assessment scores and what to do with them. In the post-intervention observation, some relevant changes were evident in the teacher's teaching and assessment practices. He adopted an integrated-skills approach, which is an important component in CA, while teaching. Next, his systematic and fair use of project-based assessment, in which he used a set of assessment criteria, was also an indication of his changing assessment practices. Unlike his pre-intervention ways, his utilization of flow-chart was a much more convincing task for assessing reading comprehension skills. A few things like instructions in mother tongue and feedback through peer- and group-correction were also significant considering that how they reflect the change in the teacher's attitude towards assessment.

During the pre-intervention interview, *Teacher 4* talked about his doubts regarding the utility of internal assessments, project-based learning and assessment of language skills and peer feedback. He also informed the researcher about his dependence on model question papers for developing question papers for internal assessment, the difficulty in

providing individual feedback and inability to convince the school administration about internal assessments. During the post-intervention interviews, however, he seemed to have found ways to counter many of his problems. He was confident that he would be able to change the system of internal assessment in his school with the help of his colleague who was a part of the intervention programme. He was sure that he would use alternative methods of assessment like portfolio, observation report and project for his next internal assessments. He also informed the researcher that he was planning to provide feedback to all his students through a set of rubrics. He showed some of these which were written on a large piece of paper. He was going to put it up on the class notice board after an internal assessment in which he would assess his students' writing skills.

6.4.2.1 Influence of Pedagogic Context and Personal Factors on Changes

Much like in the case of Case 1, the pedagogic context and the individual ability and beliefs about assessment had some impact on the changes exhibited by the Case 2 teachers in their practices. The AP State Board followed the assessment policy suggested by CCE. So integration of formative assessment with teaching, use of non-traditional assessment methods like observation, project-based and portfolio assessment, developmental grading, etc. were emphasized by the State Board. The Board policy could be a reason why the two State Board teachers were happy to use a few ideas acquired through the TD programme in their respective classrooms. Some of them included taking interest in alternative methods of assessment like observation and portfolio assessment, using a mixture of oral-written and peer-teacher feedback, making use of rubrics to provide individual feedback to students, creating and utilizing authentic tasks, etc. It can be assumed that the heavy workload at the school might have prompted them to make use of rubrics for providing feedback to individual students. But the more important thing was that the teachers thought about rubrics as a tool to save time spent on offering individual feedback. Also, the Headmaster wanted his teachers to learn more about formative assessment and had earlier asked the two teachers to share their new learning with other English teachers. This could be a reason behind the teachers' efforts to learn things that would help other English teachers as well. The teachers' individual ability and beliefs about assessment also got reflected in the form of a few changes. The third teacher's statement that it is important for teachers, students, school managements and parents to understand the meaning and importance of language proficiency showed

how the teacher understood/perceived the concept at a personal level. Other instances included the use of flow chart for assessing reading comprehension skills, instruction in mother tongue, providing every student with a chance to speak in the classroom and efforts to understand the concept of formative assessment better.

6.4.3 Case 3

Teacher 5

Interview data

The fifth teacher seemed to be totally unhappy with the change in assessment policy because she thought it would lead to unfair practice among teachers and schools and take away the seriousness among students about studies. She strongly believed that annual examinations were much better than formative assessments. However, she very strictly followed the instructions given in the teacher manual while designing and carrying out assessments in her classroom. She prepared her own question papers and deducted marks for every spelling and grammatical error irrespective of what she was assessing. She interpreted high scores as "good performance" and low scores as "poor performance". She gave individual feedback, both orally and written, only to 'poor' students. Whenever a very large percentage of her students failed in her class tests, she asked them to prepare well so that they could score better in the next one. She was convinced that she could evaluate any task by just going through it because of her experience.

Classroom observation before intervention

In the first of her observation classes, the fifth teacher taught composition to her class X students. She gave them two topics and asked them to write for 30 minutes on one of the topics. No word limit and format for writing was given to the students, which could raise doubts about her understanding of writing skills and approach to teaching the same. At the end of the class, which was of 40 minutes, she collected the answer books of the students. She informed the researcher that she would try to find as many mistakes/errors as possible from each answer script and assign low scores to all the compositions. It confirmed the doubt about her lack of understanding mentioned earlier in the paragraph. Then she would discuss the format, content, organization, and so on. She was sure that her approach would help learners work hard to write accurately and carefully. In

addition, the students would be ready to write well in their final examination. However, her examination-oriented instruction strategy was neither learner-friendly nor productive.

The teacher tried to prepare her students for an internal assessment comprising a book/film review during the second observation class. She wrote the instructions on the blackboard and talked about each one elaborately. Then she asked her students if they had any doubts. She answered all their questions related to the format, style of writing and other such relevant details. Then she warned them about spelling and grammatical errors and plagiarism and told them that they would be penalized with negative marking for those kinds of errors and if caught plagiarising. Finally, she asked them to read a book or watch a classic (movie) and then write a review. The students were given two weeks to complete the review. Though she was doing everything keeping certain objectives in mind, her entire focus was dominated by accuracy in writing. Moreover, it did not make any sense the way the composition task was handled without feedback playing any role in it.

In her third class, she conducted a classroom assessment of listening skills. The students were aware of the date and time of the assessment. The teacher wrote a set of questions on the blackboard. They included short answer type 'wh-', 'yes-no' and blank-filling questions. She asked her students to answer all the questions while listening to what she would read aloud. Then she read out a piece of news item from a newspaper to the entire class. It was a long text. After completing reading it once, she checked with the students. As some of them were not able to complete writing the answers, she read out the text once more. She asked everyone to check their answer script for grammatical and spelling errors/mistakes and warned them that they would lose marks for such errors/mistakes. On the one hand, the teacher did certain things right in the form of choosing a news item, including 'yes-no' and 'true-false' questions and reading out the text twice, on the other hand, she went off the track by selecting a lengthy text, including 'wh-' questions and focusing on grammatical and spelling errors in students' responses.

Classroom observation after intervention

In the first class that was observed after the intervention, she taught the use of prepositions to her students. It was in fact a revision class in which syllabus contents covered in the first semester were revisited. She wrote a few prepositions on the blackboard and asked students to work in pairs and make sentences using those. She

gave them 20 minutes to complete the task. After that, she asked each pair to read out the sentences for a particular preposition to the class and initiated a discussion on whether the sentences were right and whether the preposition could be used in other ways. The students were made to think and at times, she wrote a few sentences containing the use of prepositions in different contexts and drew their attention towards collocations. As homework, she asked students to take a passage from their literature book and identify how prepositions were used in them. The only noticeable change found in this class was the teacher's attitude towards her students. Making students work in pairs and organising a discussion were two learner-friendly activities.

In her next class, the teacher taught formal letter writing. Once again, it was a revision class. She asked them to write a formal letter on a topic mutually decided on with her students. It was a positive sign that she involved students in the process of decision-making. Then, she gave them 15 minutes to complete the task and a set of criteria against which their answers would be evaluated by their peers. The criteria, which were relevant to the topic, included format, adequacy and appropriateness of information, correctness and vocabulary. After fifteen minutes, she collected their scripts and made each other evaluate and score their scripts. She informed the researcher that she got the idea of peer-correction during the intervention programme. But she took the scripts home so that she could check the scripts meticulously for grammar and spelling. It was a little puzzling the way she employed peer-correction. No follow-up discussion was taken up after corrections were made by peers. Moreover, she continued focusing on grammar and spelling.

In her last class, she taught them writing informal letters. She asked her students to work in groups of 3 and write a letter to a friend in 15 minutes. They were allowed to choose the topic of their letter. Then, she made them write the purpose of their letter below the letter and exchange their work with another group. She asked each group to look at the purpose of the letter and check if the letter conveys the intended message. Moreover, they were asked to check sentence structures, use of appropriate words and salutations in the letter written by their peers. Finally, she collected the scripts and took them with her for thorough correction much like what she did in her last class observed by the researcher. Once again, peer correction fell short of being effective.

The observation of classes before the intervention revealed that *Teacher 5* misunderstood concepts like language learning, accuracy and error correction. As a result, her assessment was not learning-oriented and could also be considered as unfair. The project-based assessment carried out by her was not directed towards assessment and learning of language skills. Moreover, her rigid approach was not a good sign. Next, she tried to assess something that was not taught in her class and seemed to ignore a crucial aspect of assessment like text length during the assessment of listening skills. Some positive changes in her assessment practices were observed during the post-intervention period. The prominent ones included the regular and effective employment of peer-correction and her clear instructions for each and every classroom learning task. Apart from these, an integrated-skills approach was followed, students were involved in decisions related to choice of classroom tasks, and a set of assessment criteria was given before a task was given. There was also some improvement in terms of using appropriate context for teaching. All these factors might possibly get reflected in her classroom assessment practices.

Like some of the other teachers, *Teacher 5* openly criticized the introduction of formative tests during an interview that happened before the intervention programme. This criticism was not very evident, though not completely absent, during the post-intervention interview. Another important factor that came out during the pre-intervention interview was the overemphasis on accuracy. Again, she interpreted 'high scores' as good performers and 'poor scores' as poor performers. She reported that she had learnt to analyse and interpret these scores, and utilize the obtained data to get information about students' performance. But she continued to think that her assessments tasks would not need any revision, and her experience was enough to guide her.

Teacher 6

Interview data

The assessment practices of the sixth teacher were identical to that of the fifth teacher whom she followed as a role-model. She was aware of CCE and change in assessment policy at school level. She thought the new policy had too many loopholes and preferred summative tests over formative ones. She followed the assessment manual and carried

out assessments accordingly. Accuracy and high scores in every examination were her mantras. She said that grammatical errors in writing and speaking should not be tolerated at any cost. She wanted her students to perfect their pronunciation too. For her, low scores in tests meant lack of effort on the part of students. She provided oral and written feedback to everyone but all her feedback was corrective in nature. She took help from available question-banks while preparing question papers and showed the question paper to the fifth teacher who was also her colleague. She was sure that she did not need any further evaluation.

Classroom observation before intervention

Though she was not happy with the researcher observing her class, *Teacher 6* was not nervous when her classes were being observed. In her first class, she taught reading comprehension skills to her students. She started her class by asking students to read a piece of text from their prescribed textbook silently. After they completed reading a paragraph, she asked them questions about the tone, intention and point of view. Then, she answered the questions whenever the students had problems answering them. The questions were quite challenging for the students. In addition, she asked them the meaning of certain words used in that text. The students were asked to complete a set of reading comprehension questions as their homework. It would have been better if she had used the comprehension questions in the class and made use of a few leading questions to direct students to acquire a better understanding of tone, intonation and point of view.

In her second class, she taught them a poem. She asked each student to read out a stanza to their classmates. After that, she explained the meanings of each stanza to the class. She never asked any question to any student to see if they could understand the poem on their own. Instead, she expected her students to take notes while she was explaining the poem. It was quite clear that her main focus was not on language. So the interpretation of the poem was not based on the analysis of the language in the text. It was not surprising as the syllabus was heavily loaded with literature. Later, whenever she found someone not writing, she made it a point to go the student and asked her to be alert and keep writing. Only half of the poem could be completed in that class.

In her last class, *Teacher 6* taught debating skills to her students. She divided them into groups and gave them a topic related to role of women in Indian societies. Each group

was expected to discuss the topic in their groups before participating in the discussion. After they were ready, she asked each group to present their arguments to the class. After each group presentation, she told them about their grammatical and pronunciation errors along with the information they failed to include in their arguments. While it sounded logical when she pointed out missing information in the arguments, it did not make much sense when the teacher spent so much time talking about student's grammatical and pronunciation errors in a group discussion class. At the end of the class, she continued to give them some advice about how they could improve their accuracy in speaking.

Classroom observation after intervention

The sixth teacher also undertook revision activities in her first and second classes that were observed. She taught using reported speech in communicative contexts. It was a positive sign of change. She started the class by dividing students into pairs and asking each pair to prepare a dialogue script. In the next step, she exchanged the scripts among the pairs and asked them to report the dialogues to the class. They were given 15 minutes to complete the task. By the time they completed the task, the period got over. But the researcher observed her next class in which she continued with reported speech. Each group reported one dialogue script and the other groups corrected them whenever they found it necessary. The teacher monitored the process and gave her feedback when she found something incorrect. At the end of the class, an activity based on punctuation from a question bank was practised. The students inserted appropriate punctuations in a paragraph. The teacher made oral corrections to the errors whenever she found any. In the two classes, the teacher did not spend much time correcting students' grammar, pronunciation and spellings, which was a good sign. But she could not stop herself from doing that on occasions when the focus was not exactly accuracy.

In the third class, Teacher 6 taught "words liable to be confused". The teacher came to class with a hand-out carrying a set of sentences and these sentences contained words like 'adopt', 'adapt', 'sensitive', 'sensible', etc. She divided the class into pairs and dictated sentence-pairs to them. The students were asked to work with their partner and tell the difference between the meanings of the words that might be confused or used incorrectly because of their similarity in some way. Each pair got 5 marks for giving the correct answer and -5, for each wrong answer. It was made into a game and each pair was given equal chance to respond to the questions. Very few questions remained unanswered by

students. The teacher gave the answers to the students if all the pairs gave up. The teacher's strategy to teach using games was a learner-friendly one. This indicated a slight change in her attitude as correction was not at all in the picture.

Changes in assessment practices: observed and self-reported data

There were very few assessment–related things found during the pre- and post-intervention observation of the classes of *Teacher 6*. Her pre-intervention understanding of objectives of a reading lesson was not very problematic apart from the fact that she did not have a pre-decided focus and gave away answers while demanding only correct answers. In addition, she did not make much effort to involve the students and give them clear instructions. Moreover, she emphasised grammatical and pronunciation-related accuracy in a speaking test. She also had problems with distinguishing between language and content. Her post-intervention practices suggested that she had slightly changed her approach. She exhibited some understanding of teaching objectives and guided her students to evaluate their peers. She encouraged more student-involvement and utilized communicative tasks in the classroom. All these factors can be easily related to classroom assessments and interpreted as the teacher's efforts to improve her pedagogic practices after the intervention.

The teacher did not hide her doubts about the effectiveness of formative assessments. Her emphasis was on 'accuracy', interpretation of low score in an assessment as lack of effort and high scores, as a result of only hard work, on the part of the student. She prepared her own question papers but followed question banks to prepare those. Also, she was unwilling to accept that her assessment tasks could be evaluated. Though in the post-intervention interview, she did not speak against formative assessment, she was still not very sure about its utility. She reported that she had decided not to be very strict with grammar and pronunciation rules. She was sure that she would provide appropriate feedback to individual students on individual problems. She also informed that she had already started designing classroom assessment tasks based on the basic principles of assessment.

6.4.3.1 Influence of Pedagogic Context and Personal Factors on Changes

ICSE board had a very different assessment policy. Though 20 marks were assigned (out of the total 200) to formative assessment of language skills, the assessment components were confined to listening and speaking. Again, only four types of tasks were prescribed

for such assessments. The rigidity of the board's policy was supported by the assessment culture of the school. The school wanted its teachers to be very strict with all kinds of errors in language use and deduct marks for every single error. The Principal and the two English teachers believed that CCE was a waste of time and the Principal was unhappy with the use of English for 'communicative' purposes. She also wanted to know if the TD programme would help the two teachers in increasing student achievement in the final examination. These contextual factors were reflected in the changes found among the two ICSE teachers. Their rigidity regarding accuracy did not change much. There was little evidence to support that the teachers took any interest in CA or FA. Both of them mentioned the idea of providing feedback to their students. However, it was not very clear whether they wanted to concentrate only on correcting students' errors in the name of feedback. They also made it clear that they did not want to evaluate their assessment tasks after using them. The above data suggested that they contradicted some of their reported beliefs about assessment. There were, however, a few individual initiatives recorded in the process. Both the teachers tried to give some space to students' opinions and integrated multiple skills while teaching. While the fifth teacher tried to use a set of assessment criteria, the sixth teacher was quite aware of her own rigidity about accuracy and wanted to reduce that in future.

6.4.4 Cross-case Analysis of Changes

The data about changes in approaches to assessment practices in the classroom collected through interviews and observation of classes suggested that the three categories of teachers made use of the skills and knowledge acquired through the TD programme both in similar and dissimilar ways. It was also evident, to a certain extent, that the board assessment policy determined how teachers tried to bring changes in their practice at a personal level. There were quite a few common patterns in the changes reported by teachers and observed by the researcher. The following table presents a cross-case analysis of those changes.

Table 6.16: Cross-case Analysis of Changes

Case	Changes Specific to the Board	Common Changes
1 (CBSE)	• assessment in appropriate context	• understanding of sub-skills and teaching/assessment objectives • use of assessment criteria/rubrics • integration of skills/sub-skills • variety in questions/tasks • peer-evaluation/feedback • context of tasks • learner involvement
2 (State Board)	• project-based assessment • better understanding of 'formative tests' • portfolio assessment	
3 (ICSE)	• less emphasis on accuracy • clear instruction to students in tasks	

Though all the above mentioned changes were not found in CA contexts, they were still relevant considering that they were captured after the TD programme was completed. Research literature in the field of teacher education suggests that teachers often take a lot of time to change. However, this slow process can be tracked by keeping a record of the minor changes, which may or may not be directly related to the area focused on during the TD programme. In the current study, though the 'common changes' (Table 6.16) are directly related to CA, most of them were observed in teaching, not assessment, contexts. Considering that CA should be integrated with classroom teaching and learning, these changes were found to be relevant and important in the context of the present study.

6.5 Relationship between Teachers' CLAL and Assessment Practices

The relationship between CLAL of the teachers and their assessment practices formed the crux of the study and the main aim of this study was to explore the relationship between these two factors. In the following sub-sections, an effort is made to examine this relationship for individual teachers included under each case.

6.5.1 Case 1 CBSE

Teacher 1

Pre-intervention relationship

The analysis of the pre-intervention CLAL level and classroom assessment beliefs of the first teacher suggests a clear mismatch between her CLAL level— which was 'limited', and her reported belief about classroom assessment. Her lack of understanding of the principles of assessment, how to interpret scores, evaluate tasks and differentiate between formative and summative assessment, effective ways of providing feedback, etc., got reflected in her response to the CLAL questionnaire. But she tried to give a different picture of herself by exhibiting positive beliefs about various aspects of assessment. This mismatch could be due to the compulsion to hide her understanding of and knowledge about language assessment. Her school administration and peer pressure could be other reasons why she was trying to show herself in a positive light.

The classroom assessment practices, to a great extent, revealed her limited CLAL level. It was found that she did not spend much time in developing her assessments, lacked motivation to explore new assessment methods, almost copied questions and question patterns from old question papers, never used any assessment criteria, did not involve her students in the process of assessment and found providing individual feedback a time-consuming activity that should be avoided. There is no doubt that her classroom assessment practices were hindered by her lack of training in language assessment which was, in turn, the reason for her low level of assessment literacy. Of course, her school administration played a role in restricting her practices.

Post-intervention relationship

There was a close relationship between the CLAL and assessment practices of *Teacher 1*, as found from the analysis of the post-intervention data. The teacher responded correctly to 35 out of 38 questions in the questionnaire and she reported having strong beliefs (evident from her beliefs score of 1.82) about classroom assessment. Her overall improvement in CLAL seemed to influence her assessment practices. Apart from her overall level of tasks of 3.58, which was a big jump from 1.55, i.e., her pre-intervention overall level of tasks, she did show a great deal of improvement in her actual practices. She was found to be engaged in things like promoting learning through appropriate questions, using authentic contexts for teaching and making use of rubrics and observation notes to provide feedback to students during the post-intervention observation of her classes. Though these were not directly related to assessment practices, during the interview, she revealed that she had acquired quite a few things about assessment that she could use for teaching purposes too. She also spoke about how the intervention had helped her change her approach to assessment. She was confident about creating better assessment tasks, exploring formative methods of assessment like portfolio and observation, evaluating her own tasks and finding ways to better student achievement without going against the board and organizational policy.

Teacher 2

Pre-intervention relationship

Before the intervention, the CLAL level of *Teacher 2* was *limited.* He got only 21 questions right, which suggests that he did not have the basic ability to perform his assessment-related duties in his school. However, his score on the beliefs questionnaire was 1.73, which means he had strong and positive beliefs about classroom assessment. If his claim is to be believed, he should have had good understanding of concepts of classroom assessment. So, there is an obvious mismatch between his performance on the CLAL instrument and beliefs questionnaire. It is true that the overall task level of 1.51, which he got for the tasks he prepared before the intervention, has strong correlation with his *limited* CLAL level.

The teacher was worried about his teaching ability especially, his awareness of ELT methodology. This self-doubt affected his confidence and assessment ability negatively. In addition to this, his low CLAL level could be a reason for his apprehension about designing internal assessments, providing feedback and interpreting examination scores. He had very little knowledge about the sub-skills of language and teaching/assessment objectives. The only positive thing was that he was aware of his low assessment ability and wanted to get some training in assessment.

Post-intervention relationship

After participating in the TD programme on CA, the teacher performed very well on the CLAL instrument and reached the *adequate* level with a score of 34. This score suggests that he had adequate understanding of the principles of language assessment, knew how to plan, carry out and evaluate classroom assessment and could interpret assessment results and offer constructive feedback to students. His score 1.88 on the beliefs questionnaire corroborated this. The score indicates positive beliefs about and relatively sound understanding of CA.

The assessment practices of the teacher got enhanced along with the teacher's CLAL. The overall score for the post-intervention tasks was 3.53, which was of *good* level. Moreover, the score was consistent across tasks. Apart from the significant progress in his ability to design assessment tasks, he acquired a positive attitude towards assessment, displayed much better awareness of sub-skills of language, which was observed during his teaching of listening

using a poem and made use of peer and teacher feedback in his classroom. He wanted to try internal assessment, use rubrics and offer feedback to each student. His firmness in assertion was one of the positive impacts of the TD programme he underwent.

An Overview of the Relationship in Case 1 (CBSE)

The relationship between the CLAL and assessment practices of the Case 1 (CBSE) teachers more or less supports the hypothesis of the study. This relationship, as observed in the cases of the first and the second teacher before and after the observation, was influenced a lot by the assessment policy of the school and the board. Before the intervention, when the teachers' CLAL level was *limited*, their assessment tasks lacked focus. Important components like assessment criteria and feedback did not feature in them. However, some encouraging changes were observed in their attitude and tasks after the intervention. They made conscious efforts to utilize their newly gained knowledge and skills about CA while writing the tasks, task objectives, assessment criteria and planning to offer feedback. They also reported thinking about ways to implement the same while carrying out FAs in their respective classrooms.

6.5.2 Case 2 State Board

Teacher 3

Pre-intervention relationship

The pre-intervention CLAL of *Teacher 3* was traced through her performance on the CLAL survey instrument and beliefs questionnaire. She scored 22 and was assigned a *limited* CLAL level. It means that she had problems in understanding and applying basic principles of assessment, stating assessment objectives, developing assessment criteria, planning and carrying out assessments, providing feedback and interpreting scores. Her pre-intervention score on the beliefs questionnaire was 1.18, which was on the agree-strongly agree continuum. But there was a visible dip in the section PA in which she got a mean score of 0.5.

The assessment practices of the teacher remained below average, very much like her low CLAL level. The overall level of the tasks designed by the teacher was 1.59. All these pre-

intervention tasks and the task characteristics were assigned very low scores. During the interview and the classroom observation more evidence was found about the teacher's assessment practices. The teacher did not know how teacher-based assessment would lead to better learning and how to interpret examination/assessment scores/grades and provide appropriate feedback. She also had very little idea about alternative assessment. Her out-of-context "slip test"- found during the observation of her classes, also suggests that she was struggling with her assessment responsibilities.

Post-intervention relationship

After the intervention, the CLAL and classroom assessment practices of the teacher took a positive leap. The CLAL score crept to 32 and reached an *adequate* level. This suggests a good understanding of concepts of assessment and how to apply them in classroom situations. The overall score on the beliefs questionnaire, i. e., 1.68, was positively inclined towards "strongly agree". The achievements of the teacher in terms of CLAL, to a great extent, got reflected in her assessment practices. She was assigned 3.52 for her post-intervention assessment tasks which means, the tasks she designed were appropriate and conformed to the principles of assessment. Some positive changes were also found during the observation of her classes and interviews with her. She came very close to practicing dynamic assessment and experimented with, though not completely convinced about it, peer correction. She provided excellent feedback to individual students and tried to teach them to use feedback for avoiding further mistakes. She claimed that internal tests could be good for tracking students' progress. While displaying improved understanding of alternative methods of assessment, sub-skills of major language skills, various question-types and use of rubrics, she said that it would take time for her to clearly understand language proficiency and the meaning of test scores.

Teacher 4

Pre-intervention relationship

The teacher scored 16 on the CLAL instrument administered before the intervention. His *limited* CLAL level was an indicator of his low level of assessment literacy. In the beliefs questionnaire his overall score was 0.80, which was on the "don't know-agree" continuum.

There was some amount of self-contradiction and confusion in his responses. His lack of clarity about the concepts of classroom assessment was obvious. But he did agree that teachers need professional training in assessment to be able to carry out their assessment duties effectively.

The data collected about the assessment practices of *Teacher 4* more or less conformed to the low CLAL of the teacher. The overall level of the tasks prepared by him before the intervention was 1.24— the lowest score for any of the participating teachers. He had doubts about how formative tests would help in enhancing student achievement. He thought that formative tests put unnecessary burden on students. It was no surprise that he used paper-pencil written tests for the internals, never offered feedback on students' performance and did not want to change anything in formative tests. He had quite a few misconceptions about teaching objectives and methodology. He was sure that students' score in formative tests were indicators of their language ability and thus, improvement. He had no knowledge whatsoever of alternative methods of assessment.

Post-intervention relationship

After undergoing training during the intervention programme, the teacher scored 34 in the CLAL instrument and 1.67 on the beliefs questionnaire. The former score indicates *adequate* level of assessment literacy required to perform classroom assessment duties effectively. The latter score, which was tilted towards "strongly agree", suggests strong positive beliefs about concepts related to classroom assessment.

The improved CLAL of the teacher got reflected in his practices as evident from his *good* level ability to design classroom assessment tasks and the data collected through classroom observation and interview. The teacher prepared much more polished tasks compared to ones he did before the intervention. His integrated-skills approach to teaching, use of assessment criteria, effort to provide feedback through classroom discussions, plan to use portfolio, observation report and project work for internal assessments and attempt to provide feedback to individual students using a set of specifically-designed rubrics can be interpreted as an evolvement in his practices that might be a result of his enhanced CLAL level.

An Overview of the Relationship in Case 2 (State Board)

The relationship between teachers' CLAL and assessment practices for Case 2 was quite similar to that traced among Case 1 teachers. The State Board teachers, with their low CLAL level, found it difficult to discharge their assessment responsibilities effectively before undergoing the TD programme. They had very limited knowledge about FA, task design, use of assessment rubrics, feedback, interpretation of scores, etc. Thus, both the teachers had problems with change in assessment policy. However, positive changes in their attitude and approaches to assessment were found when they reached a better CLAL level after completing the TD programme on assessment. The improved CLAL level led to attempts by both the teachers to create better assessment tasks, utilize assessment criteria, offer feedback to students, etc. The influence of their institution on their assessment strategies was less than what the Case 1 teachers reported.

6.5.3 Case 3 ICSE

Teacher 5

Pre-intervention relationship

Much like the first four teachers, there is a link between the pre-intervention CLAL and classroom assessment practices of *Teacher 5*. She got 22 out of 38 on the CLAL instrument and 0.85 on the beliefs questionnaire. Her *limited* level of CLAL got reflected in her responses to the statements in the beliefs questionnaire. There were instances of mismatch and self-contradiction. But she believed that teachers need training in assessment.

The tasks she designed before the intervention were assigned an overall score of 1.5. Most of the tasks lacked direction and were indicative of the teacher's weak ability. This was further confirmed during the classroom observation. The listening task used by the teacher in the classroom was far from being acceptable. In addition, the teacher reported that she did not find teacher—based assessment very helpful and thought that such a policy would take away the seriousness about studies among students. Her pre-occupation with grammatical accuracy, lack of conceptual clarity about language proficiency, inadequate knowledge about the effectiveness of assessment tasks, involving students in the process of assessment and

use of assessment criteria, and belief that tests are for finding out only what students cannot do, were some of the things that confirmed her poor assessment practices.

Post-intervention relationship

Figures show that there was little change in the strong correlation between the CLAL and teacher's assessment practices after the intervention. Her scores on the CLAL instrument and the beliefs questionnaire were 35 and 1.41 respectively. They indicate an *adequate* level of CLAL combined with positive beliefs about assessment. It was noticed that she had much less confusion regarding the concepts of assessment in comparison to her pre-intervention performance.

Moreover, she reached a score of 3.84 for her post-intervention assessment tasks. In almost all the tasks, she tried to follow the principles of assessment very closely. Moreover, her performance was consistent across the tasks designed for all language skills and components. She was also found teaching grammar in context, developing a set of assessment criteria, being student-friendly in decision making and experimenting with peer correction. However, her classroom observation and interview data suggest that the changes found in her assessment tasks and her CLAL score did not really change her approach and attitude towards CA. Considering that change in attitude and approach to a great extent determines change in practices, the changes in her practices did not seem be genuine and lasting. Some contextual factors determined the course of their practice. ICSE— her school board and her school did not offer much scope for CA. Again, teaching of literature dominated English, the subject, at secondary level. For these reasons, the changes in practice did not seem to get internalized. They did not go beyond a surface level.

Teacher 6

Pre-intervention relationship

The relationship between CLAL and assessment practices of teachers, which was evident in the cases of five previous teachers, was also found in the case of *Teacher 6*. Her score on the CLAL instrument was 24, which means she had an *average* level of classroom assessment literacy. In fact, she was the only teacher among the six who participated in the study, with an *average* CLAL level before intervention. But on the beliefs questionnaire her score was

relatively lower than some of the teachers. Her score, 0.97, was very close to an overall average of 'agree'. Her responses included a few self-contradictions which indicated her want of knowledge about assessment.

In contrast to her score on the CLAL instrument, her pre-intervention tasks were assigned an average score of 1.59— a less than average score. Her inadequate knowledge about assessment could be traced in her preference for summative assessment over formative, misinterpretation of low test scores of students, an uncompromising love for accuracy, insistence on providing only corrective feedback and belief that her assessment tasks did not need any evaluation. However, she showed a better understanding of teaching methodology than other teachers though she also had problems with understanding teaching objectives.

Post-intervention relationship

The post-intervention relationship between the CLAL and classroom assessment practices of the teacher did not vary much from that of the pre-intervention. She could answer 37 questions correctly out of the total 38. It was the best score on the CLAL instrument by any teacher. She also showed some change in her beliefs about classroom assessment when she got 1.44 for her responses to the beliefs questionnaire. Overall, she displayed a very good understanding of the concepts of assessment and their application in task designing.

In line with her high CLAL score, the assessment tasks designed by the teacher after the intervention were assigned an average score of 4.02, which was of *excellent* level. The quality of her tasks, though, was not consistent across the skills and components. She was also found making use of authentic communicative contexts to teach grammar and re-thinking about her attitude towards language errors of students. However, not many changes were observed during her classroom teaching. There was little evidence to support that she had changed her attitude towards formative assessment. She continued to believe that her assessment tasks did not need any evaluation. Much like what happened in the case of the fifth teacher, the pedagogic context shaped her thinking and kept her from changing the assessment practices.

An Overview of the Relationship in Case 3

Case 3 added a different angle to the analysis of the relationship between teachers' CLAL and assessment practices. The assessment culture of the school and the ICSE assessment policy almost determined the practices of the teachers. Therefore, even though the low CLAL levels of the ICSE teachers got reflected in poor assessment practices, a higher CLAL level did not actually lead to as much change in assessment practices. In fact, the concern regarding their attitude towards and approach to CA underwent little change. The changes displayed in the design of assessment tasks, writing objectives, offering feedback, etc. in simulated settings could not be confirmed as both teachers continued to be sure that they were not going to use any of the new ideas in their respective classroom situations.

6.5.4 Cross-case Analysis of the Relationship

The relationship between CLAL and assessment practices of teachers across all the three cases did not turn out to be conclusive. Though it was possible to enhance the teachers' CLAL through the intervention programme, similar amount of change was not found in their assessment practices. The teachers under the three boards got higher CLAL scores after participating in the intervention programme with the Case 3 teachers getting higher scores than the teachers working under the two other boards. However, the change in the teachers' CA and CA-related practices was more evident and convincing in the case of CBSE and AP State Board teachers. Both these boards prescribed CCE, which provided the CBSE and State Board teachers a better platform to try out the new learning acquired through the TD programme. At a personal level, these teachers displayed much better enthusiasm and interest while attempting to change their pedagogic practices. Their attitude can be contrasted with the attitude of the ICSE teachers, who, despite designing effective assessment tasks and showing better understanding of the skills and knowledge imparted during TD sessions, made very few changes to their actual practices. What is very evident here is the role of pedagogic context in the relationship between teachers' CLAL and assessment practices. Of course, individual motivation of teachers, to some extent, can help them overcome the challenges thrown at them by the assessment policy of the board and the assessment culture in the school. But it may not be considered a way out for each and every teaching working in an environment where CA/FA is not prescribed and promoted.

6.6 Conclusion

The main purpose of the present study was to explore the relationship between CLAL and classroom assessment practices of secondary school teachers. The analysis of the data presented in the chapter suggests that there can be a positive relationship between the two aforesaid factors if the pedagogic context has a favourable attitude towards CA. Also, the data have provided evidence regarding the impact of the professional development programme in assessment on teachers' CLAL and certain important aspects of CA. In the next chapter, the hypotheses of the study are revisited, answers to the research questions are discussed, the limitations of the study are highlighted and suggestions for further research are given.

CHAPTER 7

DISCUSSION AND CONCLUSION

7.1 Introduction

The present study tried to explore the relationship between teachers' CLAL and their assessment practices. Undertaken in the state of Andhra Pradesh, the study was built on a CLAL survey of secondary school English teachers across the state. After establishing that most of these teachers possess less than adequate level of CLAL, three case studies were carried out with three pairs of English teachers working in three secondary schools— one CBSE, one Andhra Pradesh State Board and one ICSE. The relationship between their CLAL and assessment practices was recorded before and after an intervention that aimed to improve their CLAL. The analysis of the data is described in fifth and sixth chapters. In this chapter, the findings of the study are presented along with discussions on the findings, the implications of the study, suggestions for further research and limitations of the study.

7.2 Findings of the Study

Using data collected through quantitative and qualitative methods, the study tried to find answers to the research questions it proposed to address. The findings of the study are presented below in the form of answers to the research questions and follow-up discussions.

7.2.1 *What is the average CLAL level of secondary school English teachers in the state?*

According to the CLAL survey results, the average CLAL level of secondary school English teachers in the state was found to be far less than adequate, i. e., less than what they required to carry out classroom assessments they were assigned. Their knowledge about language assessment, skills to carry out classroom assessments and ability to interpret results and provide feedback were not better than what is often assumed and impressionistically claimed by educationists. Though the CLAL survey instrument was not a fool-proof tool, it provided some valuable information about teachers' assessment abilities. An effort was made to pitch the questions at their level, and the use of jargon from the field of ELT and the area of

Language Testing was also avoided in the instrument. Even then, the overall average quality of teachers' responses was not satisfactory.

Discussion

The findings from the survey only confirmed the general belief that the CLAL of secondary school English teachers is very low. It may not be very different for teachers working at primary level and in colleges. In this connection, even the language assessment literacy level of teacher educators and policy makers can be questioned. Such questions have not often been raised during academic debates and policy-related discussions and in research publications. However, it is high time the government takes steps to enhance the assessment literacy of all the stakeholders, including institutional heads and parents, in the process of assessment. Just implementing change in assessment policy is not sufficient to ensure burden-free learning. A thorough understanding of assessment at all levels can surely help the cause of learning.

7.2.2 *How is teachers' CLAL related to their assessment practices?*

The analysis of the obtained data suggested that there is a relationship between teachers' CLAL and assessment practices but it is convincing and strong only when the board and school assessment policies are in congruence with the principles of CA. In the cases of CBSE (Case 1) and State Board (Case 2) teachers, the CLAL scores got reflected in some aspects of assessment practice like writing objectives, designing better assessment tasks and criteria, providing feedback, etc and their in their beliefs as well. Even though the Case 3 (ICSE) teachers showed signs of improvement in the above mentioned aspects of assessment and beliefs, the classroom observation and interview data suggested that they did not seem to be convinced about the utility of the newly gained knowledge and skills and that they really intend to change their practice accordingly. The context of practice was the main reason behind the differences in practices between the CBSE and State Board teachers on the one hand, and ICSE teachers, on the other hand. Also, there was a gap between what the ICSE teachers claimed to believe about assessment and what they were doing as part of their assessment duties and responsibilities. Thus, the evidence was not enough to establish a 100% correspondence. These findings further confirm claims made by Mewborn (2001) that it is difficult to generalize how teachers' knowledge is reflected in their practice and the

205

sociocultural view about the importance of pedagogic context in shaping teachers' practice.

Discussion

The factors that were key to a strong relationship between teachers' CLAL and assessment practices include the assessment policy at curricular level, institutional policy about assessment, motivation of teachers to utilize assessment for pedagogic purposes and their English language ability. As the study focused only on the relationship, the impact of the above-mentioned factors were not fully explored. It was not really very clear how the CLAL of teachers who had an unfavourable institutional context increased after the intervention and why those teachers designed better tasks when asked by the researcher and brought some changes in their classroom practice. Even if the changes were found unconvincing and cosmetic in nature, there was little explanation for why teachers improved in terms of designing CA tasks and assessment criteria and plans to provide feedback and claimed change in their beliefs. Further investigation is necessary to arrive at any conclusive claim regarding the puzzling behaviour of the teachers.

It was interesting to note that there were differences between the teachers from the same school. Even though they worked in the same school and the external agencies mentioned above were same for both, their actions found dissimilar paths in terms of using training components. The intrinsic motivation of teachers along with the external factors might have shaped their actions.

Another factor that drew attention was the connection between beliefs and performance on the CLAL instrument. The findings suggested absence of uniformity. Same was also true for the link between beliefs and practices of teachers. The efforts of some of the teachers to show themselves in a positive light were obvious from their ambivalent and ambiguous responses to some of the statements in the beliefs questionnaire. In the process, they self-contradicted on occasion. Though their positive beliefs about many aspects of classroom assessment included in the questionnaire can be construed as their awareness about them, there is no certainty that the teachers had the skills to practice or practiced what they claimed to believe. However, the beliefs got stronger with the improvement in the CLAL level of the teachers. The improved knowledge about classroom assessment could have led to more

positive claims by teachers. The review of literature on teachers' beliefs about assessment (section 3.9.3) also pointed out that change in beliefs should not be overemphasized while tracing change in teachers' knowledge and practices. In this connection, Davison's (2004) suggestion that more studies on how social and pedagogic contexts shape teachers' beliefs be carried out needs some serious research attention.

7.2.3 *What impact does a short TD programme in CA of language ability have on the teachers' ability to design classroom assessment tasks and assessment criteria, provide feedback, and on the teachers' beliefs about assessment?*

The 18-hour TD programme in assessment led to some perceptible changes in the teachers' knowledge and beliefs about assessment, ability to design classroom assessment tasks and assessment criteria and provide feedback. The changes were traced in both simulated and natural contexts across all three cases. The impact did not vary significantly from one case to another. Thus, it can be claimed that the context of practice did not have much impact as far as the above mentioned factors are concerned, although the change in practices was limited to only Case 1 (CBSE) and Case 2 (State Board) teachers. The change in practice was not really evident in the case of ICSE teachers. Though the study by Chinda (2009) did not directly discuss assessment context as a crucial factor shaping the impact of TD/PD programmes on teachers, it did emphasize the importance of context-specific needs of individual teachers and asserted that the success of a programme depends a lot on the extent to which these needs are addressed in TD/PD programmes. In our context, not all the teachers had to design their internal/classroom/teacher-based assessments when the data collection for the study was happening after the intervention. Moreover, not all the three schools, in which the teachers worked, might have agreed to allow the teachers to change the institutional approach to internal assessment/formative tests. So it was almost impossible to get any conclusive data about the impact of the TD programme on their CA practices. However, analysis of the data obtained through classroom observation, evaluation of assessment tasks designed by the teachers and interview with the teachers indicated that the TD programme led to some changes in CA criteria and task design, feedback strategies and beliefs about assessment. Whether and how they put all their newly learnt skills, strategies and knowledge into practice will depend on factors like their self-motivation, support from

the school management and the assessment policy of the board under which the school functions.

The TD programme could generate some impact because of various reasons. Firstly, it was developed on the basis of the CLAL needs of the participating teachers. Secondly, it was conducted without disturbing the teachers' official schedule. Thirdly, a workshop approach was followed for the entire period of intervention. Fourthly, the relationship between the instructor and the participants was not formal and hierarchical. Thus, the findings support the claims made by Chinda (2009) and Jeong (2011) that a need-based, informal, workshop-based and integrative TD programmes can help teachers grow professionally.

Discussion

The change in the teachers' practices related to CA, as a result of their participation in the TD programme, was echoed in different aspects of practice. The common areas of development among teachers were the momentous advance in the teachers' ability to design classroom assessment tasks, write assessment objectives, develop assessment criteria and offer feedback. However, it might take a lot more time for teachers to try out what they liked and learnt during the intervention programme. They were all expected to go through a process of trial-and-error before internalizing the components. Since school administrations, to a great extent, decide the course of internal assessment, it may be tempting to evaluate the effectiveness of the programme against the achievement of teachers in convincing the school authorities and employing new approaches to assessment for their internal assessments. A year-long follow-up study can serve the purpose in this regard. Furthermore, the role of self-motivation as a guiding factor in the process of teacher learning calls for some empirical enquiry.

7.2.4 *How do teachers respond to the programme?*

The teachers' response to the programme varied from one to the other and from one week to the other. The CBSE teachers were quite enthusiastic about the programme from the start probably because they knew that it would help them carry out the internal assessments, prescribed in the form of CCE, more effectively. The State Board teachers were also excited about the programme because they thought they would learn to create better internal

assessments by participating in it. Though the motivation level was a little lower during some sessions, there was no hint that they were resistant to the contents of the programme. In both cases— CBSE and State Board— the teachers' participation was reinforced by the prescribed assessment policy, the possibility of trying out the skills and knowledge in the school and the realisation of the need to implement and utilize the policy. These factors were missing in the case of ICSE teachers.

The ICSE teachers did not have much interest in the programme. They agreed to participate and attended the sessions because the school management wanted them to. However, their stern resistance at the beginning gradually mellowed as they participated in more sessions. Their attitude got softened, and they participated in the tasks and initiated discussions after the first few weeks. However, when asked about their experience in the programme, they mentioned that they might not use what they learnt during the programme in their classroom. But their progress through the 8-weeks of training in assessment came as a surprise. They exhibited better achievement in terms of content knowledge about the aspects of CA focused in the TD programme. It is possible that they could not realize the importance of CLAL and its applications in CAs during the period traced by the study. The attempt to logically explain the significance of the TD programme for their professional growth to the ICSE teachers did not meet with much success. Even the existing literature (discussed in Chapter 3) does not provide any instances of fool-proof strategies to effect teacher change.

Discussion

The above-mentioned findings point to two interesting propositions. Since they indicate that educational and institutional policies play a major role in encouraging in-service teachers to equip themselves professionally and experiment with new ideas, it may be appropriate to immediately follow changes in policy with TD programmes catering to teachers' specific needs. The motivation and enthusiasm of the teachers from the CBSE and the State Board schools contrasted with the way the ICSE teachers kept themselves from experimenting new ideas and skills. The change in assessment policy was a major reason why the CBSE and State Board teachers wanted to learn about assessment through the TD programme.

The second proposition is related to the importance of short TD programmes. Such programmes have many advantages (section 3.7.2 of this dissertation). However, since PD of teachers should be a continuous process, short TD programmes should be provided to teachers at regular intervals. Though the current study did not focus on follow-up programmes, the need for such programmes cannot be overemphasized. If teachers' practices are tracked and recorded and follow-up programmes are designed accordingly, teachers can surely improve the quality of their pedagogic practices.

One important finding of the study was the high achievement of the ICSE teachers in terms of designing tasks and assessment criteria and planning to provide feedback. Since they had no intention of utilizing their knowledge and skills about CA in their institutional context prior to the TD programme, it is very difficult to zero down on a cause/ set of causes which led to their achievement. It could be their desire to prove and establish their potential as teachers or could be something to do with their personal beliefs. This may require a little more investigating.

7.3 Significance of the Findings

The main goal of the study was to examine the impressionistic belief that there is a connection between teachers' assessment literacy and practices. The belief was impressionistic in the sense there were efforts to train teachers in language assessment and enhance their assessment ability, but no previous researcher tried to examine whether teachers' assessment ability led to better assessment practices. Though Chinda (2009) made an attempt to look at the impact of training in developing and using rating-scales on teachers' rating practices, he was not concerned about teachers' assessment literacy per se. The current study focused on the concept of assessment literacy and to some extent, empirically established the relationship between assessment literacy and practices of teachers. By doing so, the study also questioned the universality of the concept of "language assessment literacy" and emphasized the need to analyse and include assessment duties and responsibilities of the concerned stakeholder/s and their organizational and policy-related demands before arriving at their required level of assessment literacy. Furthermore, it could lead to research and discussions on the language assessment literacy of ESL teachers in developing and under-developed countries.

Another important contribution of the study is the use of a Mixed Methods Approach in tracing the impact of the TD programme in assessment. Utilising a Mixed Methods for tracing the impact of teacher education on their teaching practices was suggested by Kubanyiova (2012). She thought such methodological innovations are necessary in the field of teacher education.

7.4 Implications of the study

The study has implications for policy makers, teacher education curriculum designers, experts in language assessment and teacher educators.

Policy makers

Only school education policy makers are referred to here. The findings of the study imply that there is a missing connection between the policy decisions and the way the decisions are conveyed to the stakeholders. The recent change in assessment policy is no less than a brilliant idea, which has potential to make learning burden-free and enjoyable. But it seems as if the effort put into conveying this message to students, parents, teachers, head masters, principals and teacher educators is not adequate. The meaning of FA is not clear to many teachers, and no one has told them what it means and how it works. Without educating parents and school-heads about it, it will be very difficult to implement the policy successfully. The sole aim of teacher-based, learning-oriented, integrated and teacher-made assessment is yet to be realized at the application level. This gap needs to be bridged, not only in this case, but also for any change in school education policy, so that the policy achieves its desired goals.

Teacher education curriculum designers

There must be some policy regarding the development of in-service programmes so that the local needs of the teachers are addressed. There should be arrangement for teachers from all kinds of schools to attend compulsory and regular TD programmes in different aspects of ELE like language teaching methodology, materials design and assessment. Considering that it may not be possible for all teachers to undergo long in-service programmes regularly, short and regular TD programmes must be developed and offered. These programmes should be designed for very small groups of teachers with similar needs. A system also needs to be

established so that in-service training sessions are followed by evaluation of teacher achievement in terms of implementing ideas and knowledge gained during the training period. Again, the evaluation data should be fed into the preparation of future programmes.

Apart from the above, some effort should go into balancing the immediate and long-term training needs of teachers while developing an in-service programme. This kind of balancing act will go a long way in strengthening teacher ability and ensuring effective implementation of the curriculum.

Once again, the importance of having a separate ELE policy making body cannot be overemphasised. It is high time an exclusive national curricular unit is formed to take care of policies, curricula, teacher education and other related areas of ELE. The country will immensely benefit from this change in the functioning system.

Experts in language assessment

This study has implications for experts in language assessment. Experts in language assessment must carry out some research and come up with national levels of LAL (language assessment literacy) meant for different stakeholders in the system of ELE. It should be a collaborative activity and followed by creation of training modules to enhance the LAL level of teachers, school-heads, teacher educators, materials writers, syllabus and curriculum designers and policy makers. They should also contribute more towards the development of training materials. The absence or scarcity of exclusive and indigenous training materials for language assessment is a serious problem. It will be interesting to see how the recent involvement in the process of global language assessment agencies like Pearson and Cambridge works.

Teacher educators

It was observed that the teachers enjoyed the training sessions because the power relationship between them and the researcher, who was also the instructor, was almost in balance. In addition, the informal approach to training and workshop nature of the sessions really helped. These aspects have implications for teacher educators. They can be very effective if they adopt an informal approach to training, value and respect teachers' experience and beliefs, empathise with them, understand their organizational set-up, listen to

their problems patiently and address these at individual level and provide them with hands-on experience during training sessions. The above-mentioned factors may sound too much to do, but if a teacher educator keeps things simple and works hard, these will be quite achievable.

7.5 Limitations of the Study

The present study was not a perfect one. It had a few limitations. The limitations which came to light in the early stages could be addressed, but very little could be done about the ones that were found midway and at the end of data collection. Some of the limitations are discussed below along with reasons as to why they could not be addressed by the researcher.

- **The CLAL survey instrument** could have been more comprehensive and included a few more aspects of classroom assessment. It could not be done for two reasons. The CLAL instrument that was created by the researcher was the first of its kind, and no such instrument existed in the existing language assessment research databases. Again, the effort to create a suitable one for secondary school English teachers in India took a long time. More time could not be spent. The second reason was that the length of the instrument was already 7-pages. Inclusion of more components would have made it difficult for teachers to respond and the return rate could have been compromised. The instrument could still be fine-tuned in terms of the items in it, and hopefully, a better version may evolve in future.

- **More number of teachers** could have been included in the state-wide CLAL survey. It could not be done because reaching teachers and convincing them to fill out the instrument was a challenging task.

- **Real assessment tasks**, which were designed by the teachers for FA purposes, could not be collected as part of the evaluation of the teachers' assessment practices. This could not be done because of several reasons. Though all the teachers agreed to share their classroom/internal assessment tasks with the researcher, three of them backed out when they were asked. All the teachers were not asked to design their internal assessments by their school authorities. So it was almost impossible to get even one assessment task for the post-intervention evaluation from everyone. The last reason was that it would have taken a year or even more to get more than one

assessment task from each teacher, had they agreed to share those. Thus, other authentic methods like classroom observation and interviews were employed to get the desired data.

- **Questions for each assessment task** designed by the teachers could have been collected. In fact, the teachers were requested to write the questions along with their proposed plan for the assessment tasks. But, all the teachers politely refused to do that. It would have been unfair to force them to write the questions as all of them had busy schedules at their respective schools. There was also the fear that they might have opted out of the intervention programme.

- **The post-intervention data** collection could have been extended to a few more months to get a better picture of the impact of the intervention programme. This could not be done because the school authorities permitted the researcher to meet the teachers and carry out the research for a particular period and not beyond that. It was not possible to find a school which would agree to a longer duration of training.

7.6 Suggestions for future research

This study, in the process of finding answers to a set of questions, has raised a few questions that can be addressed by future researchers. The first one has sprung from the CLAL survey undertaken to assess the average level of assessment literacy of secondary school English teachers in Andhra Pradesh. More such surveys can be conducted in other states and even at the national level so that a data base can be created and used for further research. Moreover, the LAL level of primary school, college and university teachers, school heads, teacher educators and curriculum designers can be traced using survey instruments of similar kind but different difficulty level. Also, the survey instrument, which was designed and used for this study, was perhaps the first of its kind in the field of language testing. The previous instruments (or questionnaires) by Plake (1993), Volante and Fazio (2007), Fulcher (2012) used for assessment literacy survey were targeted at a very different population. Whereas Plake's was related to educational assessment in general, the questionnaires by Volanted and Fazio and Fulcher only obtained self-reported data from the participants. So more research can be taken up to look into the possibility of developing context specific and skills- and ability-oriented survey instruments to capture language assessment literacy. It can be further extended to survey instruments for teacher trainers, experts, institutional heads, etc.

The second question is related to further exploration of the relationship between CLAL and assessment practices of English (language) teachers. The study supports the findings of Chinda (2009). He found that a customized rater training programme had positive impact on teachers' rating practices and beliefs. Though the current study goes a step beyond Chinda's study and focuses on two larger components like CLAL and classroom assessment practices of teachers, the exploration of the relationship between the aforesaid components can be extended to teachers in primary schools, colleges and universities. Further, these explorations can be pursued in a longitudinal manner so that more in-depth analyses can be done and the nuances of the relationship with regard to different variables can be discovered. This will add to the existing claims about the relationship.

The third question pertains to the impact of in-service teacher education programmes on teachers' pedagogical practices. It has been already proved that the success of any in-service teacher education programme depends on to what extent teachers 'internalise' (Langford, 2005) the inputs provided during the programme. The present study could not capture the process of internalization thoroughly. It calls for some serious investigation. The current study also revealed that teachers prefer to try out certain new ideas over others gained during the training. It will be interesting to look into these preferences of teachers.

The fourth question concerns how the organization or institution in which the teacher works affects the transfer of the knowledge gained through training to practice. Not many empirical studies have been undertaken to trace and explore the impact of the factor of organization on in-service teachers' practices.

The next question is about the impact of policy on teachers' response to in-service training. The current study did not explore this aspect even though it provided some data about how the teachers responded to the intervention programme. An independent study can examine the impact of policy as a variable on the quality of teachers' response to TD programmes.

The last question has to do with teacher beliefs about assessment. Though the current study focused on it, it could not provide any detailed account of teachers' beliefs about language assessment. It may need an entire study, exclusively focusing on the aforesaid topic, to throw further light on this complex area.

7.7 Conclusion

This study was a small, yet, honest attempt to explore TD in language assessment- a rarely explored topic, at least in India. It tried to argue for enhancing teachers' language assessment ability. With formative assessment gaining importance and gradually getting officially implemented in all kinds of schools in India, this study could not have been undertaken at a better time. It is hoped that the study will have a positive impact on policies and practices related to ELE, assessment of English language in schools and TD in language assessment.

BIBLIOGRAPHY

Abbuhl, R. (2011). Why, when, and how to replicate research. In A. Mackey & S. M. Gass (Eds.), *Research methods in second language acquisition: A practical guide* (pp. 296-312). London: Basil Blackwell.

Airasian, P.W. (1997). *Classroom assessment* (3rd ed.). New York: McGraw Hill.

Allport, G. (1967). Attitudes. In M. Fishbein (Ed.), *Readings in attitude theory and measurement* (pp. 1–13). New York, NY: John Wiley & Sons.

American Federation of Teachers, National Council on Measurement in Education, & National Education Association. (1990). Standards for teacher competence in educational assessment of students. *Educational Measurement: Issues and Practice,* (4), 30-32.

Anderson, L.W. (2003). *Classroom assessment: Enhancing the quality of teacher decision making.* E-Book: Mahwah, N.J. Lawrence Erlbaum Associates, Inc.

Andrews, S. (2003). Just like instant noodles: L2 teachers and their beliefs about grammar pedagogy. *Teachers and Teaching 9*(4), 351–375.

Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco: Jossey-Bass.

Arkoudis, S., & O'Loughlin, K. (2004). Tensions between validity and outcomes: Teacher assessment of written work of recently arrived immigrant ESL students. *Language Assessment Quarterly, 1*(3), 284–304.

Assessment Reform Group (2002). *Assessment for learning: 10 principles.* Cambridge: University of Cambridge School of Education.

Association of Language Testers in Europe (1998). *Multilingual glossary of language testing terms.* Cambridge: Cambridge University Press.

Avalos-Bevan, B. (2011). Teacher professional development in Teaching and Teacher Education over ten years.*Teaching and Teacher Education 27* (1), 10-20.

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language Testing, 17*(1), 1-42.

Bachman, L. (2007). What is the construct? The dialectic of abilities and contexts in defining constructs in language assessment. In J. Fox, M. Wesche, D. Bayliss, L. Cheng, C. E. Turner and C. Doe (Eds.), *Language testing reconsidered* (pp. 41-71). Ottawa: University of Ottawa Press.

Bailey, K. M. & Brown, J. D. (1996). Language testing courses: What are they? In A.

Cunning & R. Berwick (Eds.), *Validation in language testing* (pp. 236-256). Clevedon, UK: Multilingual Matters.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special?. *Journal of Teacher Education*, *59*(5), 389-407.

Ballantyne, K. G., Sanderman, A. R., & Levy, J. (2008). *Educating English language learners:Building teacher capacity*. Washington, DC: National Clearinghouse   for English Language Acquisition. Available at
http://www.ncela.gwu.edu/practice/mainstream_teachers.htm.

Bangert-Drowns, R. L., Kulik, C. C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*(2), 213–238.

Barnard, R. & Burns, A. (Eds.) (2012). Researching language teacher cognition and practice: International case studies. Bristol, UK: *Multilingual Matters.*

Barrett, A. (2010). Monitoring and evaluating INSET in India: Challenges and possible solutions. In S. Sheehan (Ed.). *Teacher Development and Education in Context: A selection of papers presented at IATEFL 2010 by the British Council* (pp. 5-16). British Council.

Bartels, N. (ed.) (2005). *Applied linguistics and language teacher education*. New York: Springer.

Bernard, R. (2000). *Social research methods*. Newbury Park, CA: Sage Publications.

Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*, 139-148.

Black, P. J., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*(1), 5-31.

Black, P. (2009). Formative assessment issues across the curriculum: The theory and the practice. *TESOL Quarterly, 43*(3), 519–524.

Bliem, C. L., & Davinroy, K. H. (1997). Teachers' beliefs about assessment and instruction in literacy. *CSE technical report* 421. Retrieved August 19, 2003, www.cse.ucla.edu/Repods/TECH421.PDF.

Bloom, B. S., Hastings, J. T. & Madaus, G. F. (1971). *Handbook of Formative and summative evaluation of student learning*. New York: McGraw-Hill.

Bolitho, R. (1986). Teacher development: A personal perspective. *Teacher Development, Newsletter of the IATEFL Teacher Development Special Group 1*, 2.

Borg, M. (2001). Teachers' beliefs. *ELTJ, 55*(2)*, 186-188*

Borg, S. (2003). Teacher cognition in language teaching: A review of research on what language teachers think, know, believe, and do. *Language Teaching, 36*(2), 81-109.

Borg, S. (2005). Teacher cognition in language teaching. In K. Johnson (Ed.), *Expertise in second language learning and teaching* (pp. 190-209). Basingstoke: Palgrave Macmillan.

Borg, S. (2006). *Teacher cognition and language education: Research and practice.* London: Continuum.

Borg, S. (2009). Language teacher cognition. In A. Burns & J. C. Richards (Eds.), *The Cambridge guide to second language teacher education* (pp. 163-171). Cambridge: Cambridge University Press.

Borg, S. (2011). Language teacher education. In J. Simpson (Ed.), *The Routledge handbook of applied linguistics* (pp. 215-228). London: Routledge.

Breen, M. P., Barratt-Pugh, C., Derewianka, B., House, H., Hudson, C., Lumley, T. & Rohl, M. (1997). *Profiling ESL children: How teachers interpret and use national and state assessment frameworks (Vols. 1-3).* Canberra: DEETYA.

Brindley, G. (1997). Assessment and the language teacher: Trends and transitions. *The Language Teacher, 21*(9), 37-39.

Brindley, G. (1998). Outcomes-based assessment and reporting in language learning programmes: A review of the issues. *Language Testing, 15* (1), 45-85.

Brindley, G. (2001). Outcomes-based assessment in practice: Some examples and emerging insights. *Language Testing, 18*(4), 393-407.

Brookhart, S. M. (2001). Successful students' formative and summative uses of assessment information. *Assessment in Education, 8*(2), 153-169.

Brookhart, S. (2010). *Formative Assessment Strategies for Every Classroom: An ASCD Action Tool.* ASCD.

Brookhart, S. (2011). Educational assessment knowledge and skills for teachers. *Educational Measurement: Issues and Practice, 30*, 3–12.

Brooks, J. G., & Brooks, M. (1993). *In search of understanding: The case for constructivist classrooms.* Alexandria, VA: ASCD.

Brown, G. T. L. (2008). Assessment literacy training and teachers' conceptions of assessment. In C. Rubie-Davies & C. Rawlinson (Eds.), *Challenging thinking about teaching and learning* (pp. 285-302). New York: Nova Science.

Brown,G.T.L. & Harris, L.R., (2010). Teachers' enacted curriculum: Understanding teacher beliefs and practices of classroom assessment. *International Association for Education Assessment Annual Conference,* Bangkok, Thailand.

Brown, H.D. (2004). *Language assessment: Principles and classroom practices.* Longman.

Brown, J. D. & Hudson, T. (1998). The alternatives in language assessment. *TESOL Quarterly, 32*(4): 653–675.

Bruner, J. S. (1975). From communication to language: A psychological perspective. *Cognition, 3,* 255–287.

Bruner, J. S. (1978). The role of dialogue in language acquisition. In A. Sinclair, R. Jarvella & W. J. M. Levelt (Eds.), *The child's conception of language.* New York: Springer-Verlag.

Bullock, D. (2010). Competency assessment for performers' list entry for those without a vocational training certificate or exemption: Final report to COPDEND. Cardiff: Cardiff University.

Bullock, D. (2011). Learner self-assessment: An investigation into teachers' beliefs. *ELT Journal, 65*(2), 114–125.

Burns, A. (1996). Starting all over again: From teaching adults to teaching beginners. In D. Freeman & J. C. Richards (Eds.), *Teacher learning in language teaching* (pp. 154−177). Cambridge: Cambridge University Press.

Busch, D. (2010). Pre-service teacher beliefs about language learning: The second language acquisition course as an agent for change. *Language Teaching Research, 14,* 318-337.

Carr, W., & Kemmis, S. (1986) *Becoming critical: Education, knowledge and action research.* Basingstoke: Falmer Press.

Cheng, L. (2005). *Changing language teaching through language testing: A washback study.* Cambridge: CUP

Cheng, L., & Wang, X. (2007). Grading, feedback and reporting in ESL/EFL classrooms. *Language Assessment Quarterly: An International Journal, 4*(1), 85-107.

Cheung, Y. L. (2013). The Impact of an In-service Professional Development Course on Writing Teacher Attitudes and Pedagogy. *Journal of Pedagogic Development, 3*(1), 11-16.

Chin, R. & Benne, K. D. (1969). General strategies for effecting change in human systems. In W. G. Bennis, K. D. Benne & R. Chin, (Eds.), *The planning of change* (pp. 32–59). New York: Holt, Rinehart & Winston.

Chinda, B. (2009). *Professional development in language testing and assessment: A case study of supporting change in assessment practice in in-service EFL teachers in Thailand* (Doctoral dissertation). University of Nottingham, UK.

Cimer, S., Cakir, I. & Cimer, A. (2010) Teachers' views on the effectiveness of in service courses on the new curriculum in Turkey. *European Journal of Teacher Education*, *33*(1), 31–41.

Cizek, G. J. (1997). Learning, achievement, and assessment: Constructs at a crossroads. In G. D. Phye (Ed.), *Handbook of classroom assessment: Learning, adjustment, and achievement* (pp. 1-33). NY: Academic Press.

Cochran-Smith, M., & Lytle, S. (1999). Relationships of knowledge and practice: Teacher learning in community. *Review of Research in Education*, *24*, 249-305.

Cohen, A.D. (1980). *Testing language ability in the classroom*. University of Michigan: Newbury House Publishers.

Cohen, A.D. (1994). *Assessing language ability in the classroom* (2nd ed.). Boston: Newbury House/Heinle & Heinle.

Conant, J. (1963). *The education of American teachers*. New York: McGraw Hill.

Cook, R. (2009). The Effects of a Short-Term Teacher Abroad Program on Teachers' Perceptions of Themselves and their Responsibilities as Global Educators. *All Graduate Theses and Dissertation. Paper 375*. Available at http://digitalcommons.usu.edu/etd/375

Cooley, V. E. (1997). Technology: Building success through teacher empowerment. *Educational Horizons, 75,* 73-77.

Coombe, C., Troudi, S. & Al-Hamly, M. (2012). Foreign and second language teacher assessment literacy: Issues, challenges, and recommendations. In C. Coombe, P. Davidson, B. O'Sullivan & P. Stoynoff (Eds.), *The Cambridge guide to second language assessment* (pp. 20-29). New York, NY: CUP.

Cox, K. B. (1994). *What "counts" in English class? A study of classroom assessment practices among California high school teachers* (Unpublished doctoral dissertation). The University of California, Riverside.

Craft, A. (2000). *Continuing professional development*. London: Routledge Falmer.

Crandall, J. A. (1994). Strategic integration: Preparing language and content teachers for linguistically and culturally diverse classrooms. In J. E. Alatis (Ed.) *Strategic*

interaction and language acquisition: Theory, practice, and research (pp. 255–274). Washington, DC: Georgetown University Press.

Creemers, B. P. M., Kyriakides, L., & Antoniou, P. (2013). Teacher professional development for improving quality in teaching. *Dordrecht, the Netherlands: Springer.*

Creswell, J. W., & Plano-Clark, V. L. (2010). *Designing and conducting mixed methods research* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Crookes, G., & Arakaki, L. (1999). Teaching idea sources and work conditions in an ESL program. *TESOL Journal, 8*(1), 15-19.

Cross, K. P., & Angelo, T. A. (1988). *Classroom assessment techniques: A handbook for faculty.* Ann Arbor, Mich.: National Center for Research to Improve Postsecondary Teaching and Learning.

Cuban, L. (1992). Curriculum stability and change. In P. Jackson (Ed.), *Handbook of research on curriculum* (pp. 216–247). New York: Macmillan.

Cumming, A. (1989). Writing expertise and second language proficiency. *Language Learning. 39*(1), 81-141.

Cumming, A. (2009). Language assessment in education: Tests, curricula, and teaching. *Annual Review of Applied Linguistics, 29*, 90–100.

Curtz, T. (2007). Teaching self-assessment. *Washington Center News, 6*(1), 23-25.

Dann, R. (2012). *Promoting assessment as learning: Improving the learning process.* Routledge.

Davies, A. ,Annie Brown, Elder, C., Hill, K., Lumley, T., & McNamara, T. (1999). *Dictionary of language testing.* Cambridge, UK: Cambridge University Press.

Davies, A. (2008). Textbook trends in teaching language testing. *Language Testing, 25*(3), 327–348.

Davison, C. (2004). The contradictory culture of classroom-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing, 20*(3), 304-333.

Davison, C. & Leung C. (2009). Current issues in English language teacher-based assessment. *TESOL Quarterly, 43*(3), 393–415.

DelliCarpini, M. (2009). Authentic assessment for ELLs in the ELA classroom. *English Journal, 98*(5), 116-119.

Desimone, L., Porter, A.C., Garet, M., Yoon, K. S., & Birman, B. (2002). Effects of professional development on teachers' instruction: Results from a three-year study. *Educational Evaluation and Policy Analysis, 24*(2), 81-112.

Dixit, K. K. (2011). *Facilitator Development for English Language Teachers' Clubs in India.* LAP Lambert Acad Publ.

Dixon-Krauss, L (1996). Vygotsky's sociohistorical perspective on learning and its application to western literacy instruction. In L. Dixon-Krauss (Ed.), *Vygotsky in the classroom: Mediated literacy instruction and assessment* (pp. 7 - 24).   Longman.

Dolan, T. (1978). Some problems in assessing language in the classroom. *British Educational Research Journal, 4*(1), 43-45.

Dörnyei, Z. (2007). *Research methods in applied linguistics.* Oxford: Oxford University Press

Dörnyei, Z., & Taguchi, T. (2010). *Questionnaires in second language research: Construction, administration, and processing* (2nd ed.). New York, NY: Routledge.

Duff, P. (*2007*). *Case study research in applied linguistics.* London: Routledge.

Duff, P. (2008). *Case study research in applied linguistics.* New York: *Lawrence* Erlbaum/Taylor & Francis.

Dunbar, S. B., Koretz, D. M., & Hoover, H. D. (1991). Quality control in the development and use of performance assessments. *Applied Measurement in Education, 4*(4), 289-303.

Earl, L.M. (2012). *Assessment as learning: Using classroom assessment to maximize student learning.* SAGE

Edelenbos, P., & Kubanek-German, A. (2004). Teacher assessment: The concept of diagnostic competence. *Language Testing, 21*(3), 259-283.

Edge, J. (1986). TEFL and teacher training – principles of integration. *English Language Research Journal, 5*, 1-9

Edge, J. (2003). Alternative discourses in teacher education. *ELT Journal,* 57(4), 386-394.

Erickson, J. A., & Anderson, J. B. (Eds.). (1997). *Learning with the community: Concepts and models for service-learning in teacher education.* Washington, DC: American Association for Higher Education.

Eswaran, S. & Singh, A. (2008). A study of effectiveness of in-service education of teachers. *All India Primary Teachers' Federation.* Retrieved from download.ei-ie. org/ .../AIPTF_India_ Effectiveness% 20of% 20Iin- Service dated 28. 12. 2011.

Falsgraf, C. (2009). The ecology of assessment. *Language Teaching, 42*(4), 491-503.

Falvey, P. & Cheng, L. *(1995). A comparative study of teachers' beliefs about assessment principles and practices. Language Testing Update, 18, 38–39.*

Farhady, H. (2003). *Classroom assessment: A plea for change.* Unpublished manuscript.

Farrell, T.S.C. (2003). Learning to teach English language during the first year: Personal issues and challenges. *Teaching and Teacher Education, 19,* 95-111.

Farrell, T.S.C (2006). Reflective practice in action: A case study of a writing teacher's reflections on practice. *TESL Canada Journal, 23*, 77-90.

Farrell, T.S.C. (2007). Learning to teach: The role of language teacher education. In G. Poedjosoedarmo, (Ed.), *Teacher education in language teaching* (pp.2-7). RELC Anthology Series 48. Singapore: SEAMEO Regional Language Centre.

Farrell, T.S.C. (2012). Reflecting on reflective practice: (Re)visiting Dewey and Schon. *TESOL Journal, 3,* 7-16.

Farrell T.S.C. (2012). Novice-service language teacher development: Bridging the gap between pre-service and in-service education and development. *TESOL Quarterly, 46*(3)*,* 435-449.

Farrell, T.S.C. (2013). *Reflective practice: Reawakening your passion for teaching.* Beijing, China: Foreign Language Teaching & Research Press.

Farrell, C. S. T., & Particia, C. P. L. (2005). Conceptions of grammar teaching: A case study of teachers' beliefs and classroom practices. *TESL-EJ Top*, *2*(9), 1-13.

Florez, M. T. & Sammons, P. (2013). *Assessment for learning: effects and impact.* Available at http://cdn.cfbt.com/~/media/cfbtcorporate/files/research/2013/r-assessment-for-learning-2013.pdf.

Frederiksen, J. R. & Collins, A. 1989. A system approach to educational testing. *Educational Researcher*, *18*(9), 27-32.

Freeman, D. (1990). Intervening in practice teaching. In J. C. Richards & D. Nunan (Eds.), *Second language teacher education* (pp. 103-117). Cambridge, UK: Cambridge University Press.

Freeman, D. (1994). Knowing into doing: Teacher education and the problem of transfer. In D. Li, D. Mahoney & J. Richards (Eds.), *Exploring second language teacher development* (pp. 1-20). Hong Kong: City University of Hong Kong.

Freeman, D. (2009). The scope of second language teacher education. In A. Burns & J. C. Richards (Eds.), *The Cambridge guide to second language teacher education* (pp. 11–19). Cambridge: Cambridge University Press.

Freeman, D. & Johnson, K. E. (2005). Towards linking teacher knowledge and student learning. In D. J. Tedick (Ed.), *Language teacher education: International perspectives on research and practice* (pp. 73-95). Mahwah, NJ: Lawrence Erlbaum Associates.

Frey, B. B. & Schmitt, V. L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academics, 18*(3), 402-423.

Frey, Bruce B., Schmitt, Vicki L., & Justin P. Allen (2012). Defining authentic classroom assessment. *Practical Assessment, Research & Evaluation*, *17*(2). Retrieved from http://pareonline.net/getvn.asp?v=17&n=2

Fuchs, L. S. & Fuchs, D. (1986). Effects of systematic formative evaluation: A meta-analysis, *Exceptional Children*, *53*, 199-208.

Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly, 9*(2), 113 - 132.

Fulcher, G. & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. New York: Routledge.

Fullan, M. G. (1991). *The new meaning of educational change*. London: Cassell Educational Limited.

Fullan, M. (2001). *The new meaning of educational change* (3rd ed.). New York: Teachers College Press.

Fullan, M. & Hargreaves, A. (1992). Teacher development and educational change. In M. Fullan & A. Hargreaves (Eds.), *Teacher development and educational change* (pp. 1–9). London: Falmer.

Garcia, G. E. & Pearson, P. D. (1994). Assessment and diversity. In L. Darling-Hammond (Ed.), *Review of research in education* (pp. 337–92). Washington, DC: American Educational Research Association.

Gebhard, J. G. (2006) *Teaching English as a foreign or second language: A self- development and methodology guide*, (2nd ed.). Ann Arber, MI: The University of Michigan Press.

Genesee, F. & Upshur, J. (1996). *Classroom-based evaluation in second language education*. Cambridge University Press.

Gipps, C. (1994). *Beyond testing: Towards a theory of educational assessment.* London: Falmer Press.

Giraldo, F. (2014). The impact of a professional development program on English language teachers' classroom performance. *PROFILE Issues in Teachers' Professional Development*, *16*(1), 63-76.

Goslin, D. A. (1967). *Teachers and testing.* New York: Russell Sage. Griffee, D. T. (2012). *An introduction to second language research methods: Design data.* Berkeley, California, USA: TESL-EJ Publications.

Grossman, P. (2005). Research on pedagogical approaches in teacher education. In M. Cochran-Smith, & K. M. Zeichner (Eds.), *Studying teacher education: The report of the AERA panel on research and teacher education* (pp. 425-476). Mahwah, NJ: Erlbaum.

Gullickson, A. R. (1993). Matching measurement instruction to classroom-based evaluation: Perceived discrepancies, needs, and challenges. *Teacher Training in Measurement and Assessment Skills.* (Paper 3). http://digitalcommons.unl.edu/burosteachertraining/3

Guskey, T. R. (2000). *Evaluating professional development.* Thousand Oaks, CA: Corwin Press.

Guskey, T. R. (March, 2002). Does it make a difference? Evaluating professional development. *Educational Leadership*, *59*(6), 45–51.

Hall, L. A. (2005). Teachers and content area reading: Attitudes, beliefs and change. *Teaching and Teacher Education*, *21*(4), 403-414.

Harvey, S, Roskvist, A., Corder, D. & Stacey, K. (2011). *An Evaluation of the Language and Culture Immersion Experiences (LCIE) for Teachers Programmes: Their impact on teachers and their contribution to effective second language learning, Report for New Zealand Ministry of Education.* Ministry of Education, New Zealand

Hamayan, E.V. (1995). Approaches to alternative assessment. *Annual Review of Applied Linguistics, 15*, 212-226.

Harbon, L. (2007). Short-Term International Experiences and Teacher Language Awareness. *International Education Journal, 8,* 229-243.

Harvey, S., Roskvist, A., Corder, D., & Stacey, K. (2011). *An Evaluation of the Language and Culture Immersion Experiences (LCIE) for Teachers Programmes: Their impact on teachers and their contribution to effective second language learning.* Wellington, New Zealand: Ministry of Education. Retrieved from http://www.educationcounts.govt.nz/

Heritage, M. (2010). *Formative assessment and next-generation assessment systems: Are we losing an opportunity?* Paper prepared for the Council of Chief State School Officers. Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from http://www.ccsso.org/Documents/2010/Formative_Assessment_Next_Generation_2010.pdf

Hesse-Biber (2010). *Mixed methods research: Merging theory with practice.* New York: Guilford Press.

Hill, K. & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing. 29*(3), 395-420.

Huerta-Macias, A. (1995). Alternative assessment: Responses to commonly asked questions. *TESOL Journal, 5*, 8-10.

Huerta-Macias, A. (2002). Alternative assessment: Responses to commonly asked questions. In J. C. Richards and W. A. Renandya (Eds.), *Methodology in language teaching: An anthology of current practice* (pp. 338-343). Cambridge University Press.

Huhta, A. (2010). *Diagnostic and formative assessment*. Oxford, UK: Blackwell.

Inbar-Lourie, O. (2008). Constructing a language assessment knowledge base: A focus on language assessment courses. *Language Testing, 25*(3), 385–402.

Jeong, H. (2011). *Perspectives of language assessment training for teachers and testing professionals* (Doctoral dissertation). University of Illinois at Urbana-Champaign.

Johnson, K. E. (1992). The relationship between teachers' beliefs and practices during literacy instruction for non-native speakers of English. *Journal of Reading Behavior, 14*(1), 83-108.

Johnson, K. E. (2009). *Second language teacher education: A sociocultural perspective.* New York: Routledge.

Johnson, K. E., & Golombek, P. R. (Eds.) (2011). *Research on second language teacher education: A sociocultural perspective on professional development.* New York: Routledge. Jupp, V. (Ed.). (2006). *The SAGE dictionary of social research methods*. London: Sage Publications.

Kane, M. (2006). Validation. In R. L. Brennan (Ed.), *Educational Measurement* (4th ed., pp. 17-64). New York: American Council on Education and Praeger Publishers.

Khonamri, F. & Salimi, M. (2010). The interplay between EFL high school teachers' beliefs and their instructional practices regarding reading strategies. *Novitas-ROYAL (Research on Youth and Language), 4*(1), 96-107.

Kolb, D. A. (1993). The process of experiential learning. In M. Thorpe, R. Edwards and A. Hanson (Eds.), *Culture and processes of adult learning: A reader* (pp. 139-155). London and New York: Routledge.

Kubanyiova, M. (2012). *Teacher development in action: Understanding language teachers' conceptual change.* London, UK: Palgrave Macmillan.

Kumaravadivelu, B. (2001). Toward a postmethod pedagogy. *TESOL Quarterly, 35*(4), 537-560.

Kurihara, Y., & Samimy, K.K. (2007). The impact of a U.S. teacher training program on teaching beliefs and practices: A case study of secondary school level Japanese teachers of English. *Japan Association for Language Teaching, 29*(1), 99-122.

Kuzborska, I. (2011). Links between teachers' beliefs and practices and research on reading. *Reading in a Foreign Language*, *23*(1), 102-128.

Lamb, M. (1995). The consequences of INSET. *ELT Journal 49*(1), 72-80.

Lamie, J. M. (2001). *Understanding change: The impact of in-service training on teachers of English in Japan.* Huntington, NY: Nova Science Publishers.

Lamie, J.M. (2002). An investigation into the process of change: The impact of in-service training on Japanese teachers of English. *Journal of In-Service Education. 28*(1), 135-162.

Lamie, J.M. (2004). Presenting a model of change. *Language Teaching Research, 8*, 115-142.

Lamie, J. M. (2005). *Evaluating change in English language teaching.* Basingstoke: Palgrave Macmillan.

Langford, P. E. (2005). *Vygotsky's developmental and educational psychology*. New York: Psychology Press.

Lantolf, J.P. (2009). Dynamic assessment: The dialectic integration of instruction and assessment. *Language Teaching, 42*(3), 355–368.

Leung, C. (2005). Classroom teacher assessment of second language development: Construct as practice. In E. Hinkel (Ed.), *Handbook of research in second language learning and teaching* (pp. 869-888). Mahwah, N.J.: Lawrence Earlbaum Associates.

Leung, C. (2007). Dynamic assessment: Assessment *for* and *as* teaching?. *Language Assessment Quarterly*, *4*(3), 257-278.

Leung, C. (2013). Classroom-based Assessment: Issues for language teacher education. In A. Kunnan (Ed.), *The companion to language assessment* (Vol. 3, pp. 1510-1519). Hoboken, NJ: Wiley-Blackwell.

Leung, C., & Teasdale, A. (1997b). What do teachers mean by speaking and listening? A contextualised study of assessment in multilingual classrooms in the English National Curriculum. In A. Huhta, V. Kohonen, L. Kurki-Suonio, & S. Louma (Eds.), *New contexts, goals and alternatives in language assessment* (pp. 291-324). Jyvaskyla: University of Jyvaskyla.

Lewy, A. (1990). Formative and summative evaluation. In Walberg, H. & Haertel, G. (Eds.), *The International Encyclopedia of Educational Evaluation*, 26-28.

Li, L. (2013). The complexity of language teachers' beliefs and practice: one EFL teacher's theories. Special issue: Second language teacher professional development: International perspectives. *The Language Learning Journal, 1*(2), 175-191.

Liu, Y. & Fisher, L. (2006). The development patterns of modern foreign language student teachers' conceptions of self and their explanations about change: Three cases. *Teacher Development 10*(3), 343-360.

Liu, S., Wang, J., Nam, Y., Bhattacharya, D., Karahan, E., Varma, K. & Roehrig, G. (2012). In-service teachers' attitudes and beliefs about climate change. *Paper presented at the Annual Meeting of the Association for Science Teacher education*, Clear Water Beach, Florida, January 4 – January 7.

Llosa, L. (2007). Validating a standards-based classroom assessment of English proficiency: A multitrait-multimethod approach. *Language Testing, 24*(4), 489-515.

Llosa, L. (2011). Standards-based classroom assessments of English proficiency: A review of issues, current developments, and future directions for research. *Language Testing, 28*(3), 367-382.

Lorduy,D., Lambraño, E., Garcés, G. & Bejarano, N. (December 2009). In-service English teacher's beliefs about culture and language methodology: An exploratory research in Montería. University of North Colombia: *Zona Próxima, 11*, 32-51. Available at http://www.redalyc.org/articulo.oa?id=85313003003

Luchini, P. L. (February, 2010). The importance of incorporating INSET programs into second language instruction for the attainment of professional development. *MJAL* (Special Issue)*, 2*(2). Retrieved from http://www.mjal.org/Journal/importance_incorporating_INSET.pdf

Mackey, A., & Gass, S. M. (2005). *Second language research: Methodology and design.* Mahwah, NJ: Lawrence Erlbaum Associates.

Malone, M.E. (2007). Oral proficiency assessment: The use of technology in test development and rater training. *CAL Digest*. Retrieved from http://www.cal.org/resources/digest/oralprof.html

Malone M. E. (2008). Training in language assessment. In E. Shohamy (Ed.), *Language testing and assessment*, *Vol. 7, Encyclopedia of language and education* (pp. 225-239). New York: Springer.

Malone, M.E. (2011). Assessment literacy for language educators. *CAL Digest*. Retrieved from http://www.cal.org/resources/digest/assessment-literacy-for-language-educators.html

Mann, S. (2005). The language teacher's development. *Language Teaching*, *38*(3), 103-108.

Mann, S. (2008). Teachers' use of metaphor in making sense of the first year of teaching. In T. S. C. Farrell (Ed.), *Novice language teachers* (pp. 11-28). London: Equinox.

Mata, L. (2012). Key factors of curriculum innovation in language teacher education. *World Academy of Science, Engineering and Technology*, *66,* 512-520.

Mathew, R. (2006). Tracing the After-life of Teacher Development Programmes: Reopening Closed Chapters. *English Language Teacher Education and Development. 9*(Winter), 21-38.

Mathew, R. (2012). Linking assessment to learning: what does it mean?. *Teacher Plus, May-June.* Available at http://www.teacherplus.org/2012/may-june-2012/linking-assessment-to-learning-what-does-it-mean

Mathew, R. (2014). *Teacher Development as the Future of Teacher Education.* Paper presented at the International English Language Teacher Educator Conference, Hyderabad, India.

Mathew, R. & Poehner, M. E. (2013). Monitoring Progress in the Classroom.In A. J. Kunnan (Ed.), *The Companion to Language Assessment. Vol. II: Approaches and Development, Part 6: Assessment and Learning* (pp. 631–645). Hoboken, NJ: Wiley and Sons.

Mayo, S. T. (1964). What experts think teachers ought to know about educational measurement. *Journal of Educational Measurement*, *1*(1), 79-86.

Marzano, R.J. (2006). *Classroom assessment & grading that work*. Alexandria, VA: Association for Supervision and Curriculum Development.

McGrath, I. (1986). The balancing act: Our major responsibility. *Teacher Development: Newsletter of the IATEFL Teacher Development Special Group 4,* 1.

McKay, S. L. (2006). *Researching second language classrooms*. Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

McKay, P. (2006). *Assessing young language learners.* Cambridge: Cambridge University Press.

McKay, P. (2007). The standards movement and ELT for school-aged learners: Cross-national perspectives. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching* (pp. 439-456). New York, NY: Springer

McMillan, J. H. (2003). Understanding and improving teachers' classroom assessment decision making: Implications for theory and practice. *Educational Measurement: Issues and Practices*, *22*(4), 34-44.

McNamara, T. (2006). Validity and values: Inferences and generalizability in language testing. In M. Chalhoub-Deville, C. A. Chapelle & P. A. Duff (Eds.), *Inference and generalizability in applied linguistics: Multiple research perspectives* (pp. 27-45). Amsterdam: John Benjamins.

McNamara, T. & Roever, C. (2006). *Language testing: The social dimension.* Oxford: Blackwell.

Mehrens, W. A. (1992). Using performance assessment for accountability purposes. *Educational Measurement: Issues and Practice, 11*(1), 3-9, 20.

Mertler, C.A. (2002). *Classroom assessment literacy inventory.* Adapted from the *Teacher assessment literacy questionnaire* (1993), by Barbara S. Plake & James C. Impara, University of Nebraska-Lincoln, in cooperation with The National Council on Measurement in Education & the W. K. Kellogg Foundation.

Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessment. *Educational Researcher, 23*(2), 13-23.

Mewborn, D. (2001). Teachers content knowledge, teacher education, and their effects on the preparation of elementary teachers in the United States. *Mathematics Teacher Education and Development*, *3*, 28-36.

Miles, M.B. (Ed.). (1964). *Innovation in education.* Columbia Teachers' College, New York: Bureau of Publications.

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Newbury Park, CA: Sage Publications, Inc.

Mills, A. J., Durepos, G., & Wiebe, E. (2010). *Encyclopedia of case study research (Vol. One).* California: Sage.

Ministry of Human Resource Development (2009). *Teachers in position in primary schools, Andhra Pradesh survey reports.* Retrieved from http://aises.nic.in/downloadFlash/PS/28/PS68_Secondary_Total.pdf

Mohanty, S.P. (2014). In-Service Training spells its Impact on Classroom Practices: Observation of Some Significant Variables at Elementary School Level. Scholarly Research Journal for Interdisciplinary Studies (SRJIS) *II*(XI), 114-123.

Mok, J. (2011). A case study of students' perceptions of peer assessment in Hong Kong. *ELT Journal, 65*(3), 230–239.

Moody, M.S. & Stricker, J.M. (2009). *Strategic design for student achievement.* Teachers College Press.

Mousavi, A. (2002). *An encyclopedic dictionary of language testing* (3rd ed.).   Taiwan: Tung Hua Book Company.

Muñoz, A., Palacio, M. & Escobar, L. (2012). Teachers' beliefs about assessment in   an EFL context in Colombia. *Profile*, *14*(1), 143-158.

NCERT. (2005). *National curriculum framework 2005.* New Delhi: NCERT.

NCERT. (2006). *Position paper, national focus group on examination reforms.* New Delhi: NCERT.

NCERT. (2006). *Position paper, national focus group on curriculum, syllabus and textbooks.* New Delhi: NCERT.

NCERT. (2006). *Position paper, national focus group on teaching of English.* New Delhi: NCERT.

NCFTE. (2010). *National curriculum framework for teacher education: Towards preparing professional and humane teacher.* New Delhi: NCTE.

NCTQ. (2012). *National centre on teacher quality: Yearbook brief: state of the states.* Available at http://www.nctq.org/dmsStage/State_of_the_States_2012_Teacher_Effectiveness_Policies_NCTQ_Report

Newmann, F. M. (1991). Linking restructuring to authentic student achievement. *Phi Delta Kappan, 72*(6), 458–463.

Niu, R. and Andrews, S. (2012). Commonalities and discrepancies in L2 teachers' beliefs and practices about vocabulary pedagogy: A small culture perspective. *TESOL Journal*, *6*, 134-154.

O'Loughlin, K. (2006). Learning about second language assessment: Insights from a postgraduate on-line subject forum. *University of Sydney Papers in TESOL, 1*, 71 – 85.

O'Malley, J. M., & Pierce, L. V. (1996). *Authentic assessment for English language   learners: Practical approaches for language teachers.* Reading, MA: Addison-Wesley.

O'Malley, J. M. & Pierce, L.V. (2002). *Authentic assessment for English language learners: Practical approaches for teachers* (2nd ed.). London: Longman.

O'Sullivan, B. (Ed.) (2010). *Language testing: Theory and practice*. Oxford: Palgrave.

Obeg, C. (2009). Guiding classroom instruction through performance assessment. *Journal of Aviation Management & Education, 1,* 1.

Ohata, K. (2007). Teacher development or training? Recent developments in second/foreign language teacher education. *The language research bulletin, 22*(1), 1-16. Retrieved from http://web.icu.ac.jp/lrb/vol_22/Kota%20LRB%20V22.pdf

Onwuegbuzie, A. J., & Collins, K. M. T. (2007). A typology of mixed methods sampling designs in social science research. *The Qualitative Report*, *12*(2), 281-316.

Opfer, V. D., Pedder, D. & Lavicza, Z. (2011). The Role of Teachers' Orientation to Learning in Professional Development and Change: A National Study of Teachers. England.*Teaching and Teacher Education, 27*(2), 443 – 453.

Pacek, D. (1996). Lessons to be learnt from negative evaluation. *ELT Journal*, *50*(4), 335-343.

Pachler, N. & Redondo, A. (2014). *A Practical Guide to Teaching Foreign Languages in the Secondary School (2nd ed.).* Routledge.

Padwad, A. (2011). Professional development- Why bother?. In P. Gunashekar, A. Padwad and D. Pawelec (Eds.), *Starting, stimulating and sustaining English language teacher education and development* (pp. 10-15). British Council.

Padwad, A. & Dixit, K. K. (2008). Impact of professional learning community participation on teachers' thinking about classroom practices. *TESL-EJ 12*(3). Available at http://www.tesl-ej.org/wordpress/issues/volume12/ej47/ej47a10/

Padwad, A. & Dixit, K. K. (2014). Exploring Continuing Professional Development: English Teachers' Clubs in Central India. In T. Wright and M. Beaumont (Eds.), *Experiences of Second Language Teacher Education.* (pp. 153–174). Palgrave Connect. Retrieved from

http://www.palgraveconnect.com/pc/doifinder/10.1057/9781137316257.0013

Pajares, M.F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research. 62*(3), 307-332.

Perie, M., Marion, S. F., & Gong, B. (2007). *Moving Towards a Comprehensive Assessment System: A Framework for Considering Interim Assessments.* Dover, NH: The National

Center for the Improvement of Educational Assessment, Inc. Available at: www.nciea.org.

Pawlak, M. (Ed.) (2011). *Extending the boundaries of research on second language learning and teaching.* New York: Springer.

Phipps, S (2007). What difference does DELTA make? *Research Notes*, *29*, 12-15.

Plake, B. S., & Impara, J. C. (1992). *Teacher competencies questionnaire description.* Lincoln, NE: University of Nebraska.

Poehner, M.E. (2009). Group dynamic assessment: Mediation for the L2 classroom. *TESOL Quarterly, 43*(3), 471–491.

Popham, W. J. (1995). *Classroom assessment: What teachers need to know.* Needham Heights, MA: Allyn and Bacon.

Popham, W. J. (1999). Where large scale assessment is heading and why it shouldn't. *Educational Measurement: Issues and Practice, 18*(3), 13–17.

Popham, W. J. (2004). Why assessment illiteracy is professional suicide. *Educational Leadership*, *62*(1), 82-83.

Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4–11.

Pryor, J. & Crossouard, B. (2008). A sociocultural theorisation of formative assessment. *Oxford Review of Education 34*(1), 1–20

Purpura, J. E. (2004). *Assessing grammar.* New York: CUP.

Qian, D. (2008). English language assessment in Hong Kong: A survey of practices, developments and issues. *Language Testing, 25*(1), 85-110.

Raina, M. S. P. (2005). *Opinion of secondary school teachers on the effectiveness of in-service training in enhancing professional competency.* (Unpublished M.Ed. Dissertation), University of Calicut, Calicut. Retrieved from http:// shodhganga. inflibnet. ac.in/ bitstream/10603/1546/18/18 _chapter %202.pdf.

Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing, 18*(4), 429-462.

Rea-Dickins, P. (2004). Understanding teachers as agents of assessment. *Language Testing, 21*(3), 249-258.

Rea-Dickins, P. (2007). Classroom-based assessment: Possibilites and pitfalls. In C. Davison & J. Cummins (Eds.), *The international handbook of English language teaching* (pp. 505-520). Norwell, MA: Springer Publications.

Rea-Dickins, P. (2008). Classroom-based language assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Encyclopedia of language and education* (2nd ed.), *Volume 7: Language testing and assessment* (pp. 257-271). Heidelberg, Germany: Springer.

Richards, J. C. (1990). The teacher as self-observer. In J. C. Richards (Ed.), *The language teaching matrix* (pp. 118-143). New York: Cambridge University Press.

Richards, J. C. (1996). Teachers' maxims in language teaching. *TESOL Quarterly, 30*, 281–296.

Richards, J. C. (2008). Second language teacher education today. *RELC Journal, 39*, 158-177.

Richards, J. C. & Farrell, T. S. C. (2005). *Professional development for language teachers.* Cambridge: Cambridge University Press.

Richards, J. C., Gallo, P. B. & Renandya, W. A. (2001). Exploring teachers' beliefs and the processes of change. *PAC Journal, 1*(1), 41-58.

Richards, J. C., & Nunan, D. (Eds.). (1990). *Second language teacher education.* Cambridge, UK: Cambridge University Press.

Richardson, V. (1996). The role of attitudes and beliefs in learning to teach. In J.Sikula (Ed.), *Handbook of research on teacher education* (2nd ed., pp. 102-119). New York: Macmillan.

Richardson, V. & Placier, P. (2001). Teacher change. In V. Richardson (Ed.) *Handbook of research on teaching* (4th ed., pp. 905-947). Washington, DC: American Educational Research Association.

Rodriguez, M.C. (2004). The Role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education, 17*(1), 1–24.

Rodrigues, S. (2005a) *Model of Teacher Professional Development: The Partnership in Primary Science Project.* Nova Science Publishers, Inc: Hauppauge.

Rodrigues, S. (2005b) Teacher Professional Development in Science Education. In *International Perspectives on Teacher Professional Development*, Rodrigues, S. (editor) Nova Science Publishers, Inc. 1- 13.

Rogier, D. (2009). *English language teachers' perspectives on assessment.* In partial fulfillment of EED 2009 Language Teacher Education module, Exeter, UK: The University of Exeter.

Ruddock, J. (1981) Making the Most of the Short In–service Course School Council. *Working Paper 71.* London: Methuen.

Russell, M. K., & Airasian, P. W. (2012). *Classroom assessment concepts and applications* (7th ed.). New York: McGraw-Hill.

Savignon, S. J., & Wang, C. (2003). Communicative language teaching in EFL contexts: Learner attitudes and perceptions. *IRAL, 41*(3): 223–49.

Schafer, W. D. (1993). Assessment literacy for teachers. *Theory into practice*, *32*(2), 118-126.

Scholz, R. W., & Tietje, O. (2002). Embedded case study methods: Integrating quantitative and qualitative knowledge. London: Sage.

Scott, R. & Rodgers, B. (1995). Changing teachers' conceptions of teaching writing: A collaborative study. *Foreign Language Annals, 28*, 234-246.

Sen, J. (2002). *Understanding Teachers' Professional Growth Through their Articulation of Personal Practical Knowledge.* (Unpublished doctoral dissertation). .English and Foreign Languages University, Hyderabad.

Seong, Y. (2011). The teacher's role in classroom-based language assessment. Teachers College, Columbia University *Working Papers in TESOL & Applied Linguistics*, *11*(1), 50-52.

Shavelson, R.J. (2006). On the Integration of Formative Assessment in Teaching and Learning: Implications for New Pathways in Teacher Education. In F. Oser, F. Achtenhagen, & U. Renold (Eds.), *Competence-Oriented Teacher Training: Old Research Demands and New Pathways.* Utrecht, The Netherlands: Sense Publishers.

Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., 1066–1101). Washington, DC: AERA.

Shermis, M. D., & Di Vesta, F. J. (2011). *Classroom assessment in action.* Lanham, MD: Rowman & Littlefied.

Shepard, L. A. (2000). The role of assessment in a learning culture. *Educational Researcher, 29*(7), 4–14.

Shohamy, E. (1998). Inside the "black box" of language tests. *Studia Anglica Posnaniensia, XXXIII*, 343-352.

Shohamy, E. (2008). Introduction to Volume 7: Language testing and assessment. In E. Shohamy & N. H. Hornberger (Eds.), *Language testing and assessment: Encyclopedia of language and education* (Vol. 7, pp. xiii–xxii). New York: Springer.

Short, D. (1993). Assessing integrated language and content instruction. *TESOL Quarterly. 27*.627—656.

Shulman, L. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher*, *15*(2), 4-14.

Sim, J. Y. (2011) *The impact of in-service teacher training: A case study of teachers' classroom practice and perception change* (Unpublished doctoral dissertation). University of Warwick: UK.

Singh, V. K. (2011). An Analytical Study of English Language Teachers' Training in India. In P. Gunashekar, A. Padwad & D. Pawelec (Eds.), *Starting, Stimulating and Sustaining English Language Teacher Education and Development* (pp. 128-135). British Council.

Smith, J. K., Smith, L. F., & De Lisi, R. (2001). *Natural classroom assessment.* Thousand Oaks, CA: Corwin Press (Sage Publications).

Smith, C., Hofer, J., Gillespie, M., Solomon, M., & Rowe, K. (2003). *How teachers change: A study of professional development in adult education.* (Report No. 25a). Cambridge, MA: National Center for the Study of Adult Learning and Literacy.

Spolsky, B. (2008). Introduction: Language testing at 25: Maturity and responsibility? *Language Testing, 25*(3), 297–305.

Steadman, M. (1998) Using classroom assessment to change both learning and teaching, *New Directions for Teaching and Learning, 75*, 23–35.

Stiggins, R. (1991). Relevant classroom assessment training for teachers. *Educational Measurement: Issues and Practice, 10*(1), 7-12.

Stiggins, R. (1993). Two disciplines of educational assessment. *Measurement and Evaluation in Counselling and Development, 26*, 93-104.

Stiggins, R. (1999). Evaluating classroom assessment training in teacher education programs. *Educational Measurement: Issues and Practice*, *18*(1), 23–27.

Stiggins, R. (2002). Assessment crisis: The absence of assessment for learning. *Phi Delta Kappan*, *83*(10), 758-65.

Stiggins, R. & Conklin, N. F. (1992). *In teachers' hands: Investigating the practices of classroom assessment.* Albany: State University of New York Press.

Stiggins, R. (2004). *Student-involved assessment for learning (*4ᵗʰ ed.*).* Upper Saddle River, NJ: Prentice-Hall, Inc.

Stiggins, R. (2007). Conquering the formative assessment frontier. In J. McMillan (Ed.), *Formative Classroom Assessment.* NY: Columbia University Teachers College Press.

Stiggins, R., Arter, J., Chappuis, J., & Chappuis, S. (2004). *Classroom assessment* for *student learning: Doing it right—using it well.* Portland, OR: Assessment   Training Institute.

Stivers, J. & Cramer, S. F. (2009). *A teacher's guide to change: Understanding,   navigating, and leading the process.* Thousand Oaks, CA: Corwin Press.

Stoynoff, S. J.   (2007). Assessing communicative competence: From theory to practice. In J. Lui (Ed.), *English language teaching in China* (pp. 127-149).   London: Continuum.

Stoynoff, S. (2009). Recent developments in language assessment and the case of four large-scale tests of ESOL ability. *Language Teaching, 42*(1), 1–40.

Stoynoff, S. (2012), Looking backward and forward at classroom-based language assessment, *ELT Journal, 66*(4), 523-532.

Stoynoff, S. & Coombe, C. (2012). Professional development in language   assessment. In C. Coombe, P. Davidson, B. O'Sullivan & P. Stoynoff (Eds.), *The Cambridge guide to second language assessment* (122-129). New York,   NY: CUP.

Subrahmanian, T. (2001). *The impact of DIETs on the work efficiency of primary school teachers of Kerala State.* (Doctoral Thesis), Department of Education. Calicut University.

Tatto, M. (1998). The influence of teacher education on teachers' beliefs about purposes of education, roles, and practice. *Journal of Teacher Education, 49*(1),   66-77.

Tavakoli, H. (2012). *A dictionary of research methodology and statistics in applied linguistics.* Tehran: Rahmana Press.

Taylor, L. (2009). Developing assessment literacy. *Annual Review of Applied Linguistics, 29*, 21–36.

Teachers of English to Speakers of Other Languages, Inc. (1998). *Managing the assessment process: A framework for measuring student attainment of the ESL standards.* Alexandria, VA: Author.

Tedick, D. (Ed.). (2005). *Second language teacher education: International perspectives.* Mahwah, NJ: Erlbaum.

Theodoropulos, C. (2011). Classroom-based language assessment and L2 learning. Teachers College, Columbia University. *Working Papers in TESOL & Applied Linguistics, 11*(1).

Theriot, S., & Tice, K. C. (2009). Teachers' knowledge development and change: Untangling beliefs and practices. *Literacy Research and Instruction, 48*, 65-75.

Tierney, R. D. (2006). Changing practices: Influences on classroom assessment. *Assessment in Education*, *13*(3), 239-264.

Tomlinson, B. (1988). In-Service TEFL, Is it worth the risk? *The Teacher Trainer, 2*(2), 17-19.

Torrie, M. C., & Van Buren, J. B. (2008). Student and program assessment: Assessment literacy, the basis for student assessment. *Journal of Family and Consumer Sciences Education, 26*(National Teacher Standards 3), 36-49. Available at http://www.natefacs.org/JFCSE/v26Standards3/v26Standards3Torrie.pdf

Troudi, S., Coombe, C. & Al-Hamly, M. (2009). EFL teachers' views of English language assessment in higher education in the United Arab Emirates and Kuwait. *Tesol Quartlery*, *43*(3), 546-555.

Tsui, A. B. M. (2011). Teacher education and teacher development. In E. Hinkel (Ed.), *Handbook of research in second language teaching and learning* (Vol. II, pp. 21-39). New York & London: Routledge.

Ur, P. (1997). The English teacher as professional. *English Teaching Professional 1*(2), 3-5.

Veenman, S., Tulder, M. V. & Voeten, M. (1994). The impact of in-service training on teacher behaviour. *Teaching and Teacher Education 10*(3), 303-17.

Vefalı, G.M. & Tuncergil, Ç. (Winter, 2012). Exploring in-service English language teacher trainees' and trainers' practice and beliefs in Northern Cyprus. *ELTED, 15*.

Vogt, K., Guerin, E., Sahinkarakas, Tsagari, D., S., Pavlou, P., & Affiri, Q. (2008). Assessment literacy of foreign language teachers in Europe – Current trends and future perspectives. *Poster presented at the 5th Conference, Athens, Greece, May 8th-11, 2008*. Available at www.ealta.eu.org/conference/2008/programme.htm

Volante, L., & Fazio, X. (2007). Exploring teacher candidates' assessment literacy: Implications for teacher education reform and professional development. *Canadian Journal of Education, 30*(3), 749–770.

Vygotsky, L. S. (1978). Interaction between learning and development (M. Lopez Morillas, Trans.). In M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.), *Mind in society: The development of higher psychological processes* (pp. 79-91). Cambridge, MA: Harvard University Press.

Wallace, M. J. (1991). *Training foreign language teachers: A reflective approach.* Cambridge: Cambridge University Press.

Waters, A. (2006). Facilitating follow-up in ELT INSET. *Language Teaching Research, 10*(1), 32-52.

Waters, A. & Vilches, M. L. C. (2000). Integrating Teacher Learning: the School-Based Follow-up Development Activity. *ELT Journal, 54*(2), 126-134.

White, E. (January 11, 2009). *Assessment literacy for effective classroom-based assessment.* Available at http://www.slideshare.net/ewhite/assessment-literacy-effective-classroombased-assessment-presentation-907377

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan, 79*(7), 703-713.

Wiliam, D., Lee, C., Harrison, C., & Black, P. (2004). Teachers developing assessment for learning: Impact on student achievement. *Assessment in Education, 11*, 49-65.

Wolter, B. (2000). A participant-centered approach to INSET course design. *ELT Journal, 54*(4), 311-318.

Woods, D. (1996). *Teacher cognition in language teaching: Beliefs, decision-making and classroom practice.* Cambridge, England: Cambridge University Press.

Wright, N. (2010). Twittering in teacher education: Reflecting on practicum experiences. *Open Learning: The Journal of Open, Distance and e-Learning, 25*(3), 259–265.

Yadav, S. K. (2012). *Impact of in-service teacher training on classroom transaction. (A set of four books).* New Delhi: NCERT.

Yan, C. (Winter 2007). Teachers' needs: An important factor for longer- term sustainability of cross-cultural inset programmes. *ELTED, 10.*

Yang, T. (2007). *Factors affecting EFL teachers' classroom assessment practices of young language learners.* The University of Iowa: ProQuest.

Yin, R. K. (2003). *Case study research: Design and methods* (3rd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Yin, M. (2010). Understanding classroom language assessment through teacher thinking research. *Language Assessment Quarterly, 7*(2), 175-194.

Zeichner, K., & Gore, J. (1990). Teacher socialization. In W. R. Houston (Ed.), *Handbook of research on teacher education*, (pp. 329-348). New York: Macmillan.

List of Appendices

APPENDIX A

CLAL Survey Sampling

Research Randomizer Results:
1 Set of 120 Unique Numbers Per Set
Range: From 1 to 198812 -- Sorted
Set 1

Number	Region	#
152		1
258		2
3460		3
3709	Adilabad	4
4385		1
5683		2
7491		3
7981		4
9268	Anantapur	1
10469		2
13336		3
13899		4
15041		1
16308	Chittoor	2
22031		3
28667		4
29228		1
31244	Cuddapah	2
32828		3
34377		4
34616		1
36110	East Godav	2
37775		3
41472		4
43514		1
45014		2
45498		3
46915	Guntur	4
49187		1
49473		2
49659		3
50575		4
53344		1
53663		2
53914		3
54474		4
55949		1
56683	Hyderabad	2
57536		3
59143		4
61569		1
61845		2
62734		3
64373		4
65479		1
73244		2

Number	Region	#
73373	Karimnagar	3
75557		4
76402		1
77058		2
77830		3
78655	Khammam	4
80932		1
81169		2
81261		3
84448		4
86378		1
86457		2
87637	Krishna	3
89378		4
89830		1
90251		2
92377		3
94154		4
96534	Kurnool	1
96554		2
97907		3
100144		4
100224		1
101533		2
104945		3
105487	Mahbubna	4
106131		1
107193		2
109514		3
109731		4
110443		1
110698	Medak	2
111074		3
112682		4
115712		1
116171		2
116408		3
117691	Nalgonda	4
120571		1
120662		2
121883		3
125995		4
126600		1
127493	Nizamabad	2
128303		3
129280		4
131335		1
133861	Prakasam	2
139752		3
140918		4

Number	Region	#
143868		1
144418		2
145402	Rangareddi	3
149793		4
150894		1
152436		2
152614		3
153715		4
157967	Srikakulam	1
163895	Nellore	2
167671		3
169423	Visakhapat	4
172795		1
176863		2
178039	Vizianagara	3
178663		4
181343		1
183663		2
184285	Warangal	3
184355		4
184513		1
188212		2
197399	West Goda	3
197821		4

APPENDIX B

Classroom Assessment Literacy Survey

Dear teacher,

I thank you for agreeing to participate in this survey. The aim of this survey is to measure the classroom assessment literacy of secondary school English teachers in Andhra Pradesh.

Read the questions carefully before you choose your answer from the choices given. Please contact me immediately at 09160796502 (mobile phone) or <santosheflu@gmail.com> if you have any doubt about or difficulty in understanding any of the following questions. I assure you that your responses will be used only for research purposes and your name and other details will not be disclosed. Thanks a lot!

Best regards

Santosh Mahapatra, Ph. D. (English) Student, University of Hyderabad

Mobile No. 9160796502

Your name:

Name of the school and place where you work:

Age:

i) 20-30   ii) 30-40   iii) 40-50   iv) More than 50 years

Sex:

Teaching experience:

i) Less than 5 years   ii) 5-10 years

iii)10-15 years   iv) More than 15 years

1. Which of the following options covers maximum number of purposes of a language assessment?

a) Finding out students' problems with learning, ability to use certain skills and components, overall proficiency in a language, mastery of the language skills and content taught in a prescribed syllabus, ability and readiness to be part of a particular language programme and ability to remember the texts taught in the classroom [ ]

b) Finding out students' problems with learning, ability to use certain skills and components, overall proficiency in a language, mastery of the language skills and content taught in a prescribed syllabus [ ]

c) Finding out students' problems with learning, ability to use certain skills and components, overall proficiency in a language, mastery of the language skills and content taught in a prescribed syllabus and evaluating the effectiveness of a language programme [ ]

d) Finding out students' problems with learning, ability to use certain skills and components, overall proficiency in a language, mastery of the language skills and content taught in a prescribed syllabus, ability and readiness to be part of a particular language programme and evaluating the effectiveness of a language programme [ ]

2. Three teachers in a school assess their students' oral proficiency. The first one asks her students to interact with a person, the second one asks her students to talk about a situation presented in a picture, whereas, the third one directs her students to write a set of dialogues that two persons may use to talk about their plans for summer vacation.

In the above context, the three teachers use three different:

a) Methods of assessment [ ]   b) Criteria of assessment [ ]

c) Principles of assessment [ ]   d) Approaches to assessment [ ]

3. In which of the following instances, the principle of reliability is violated?

a) An assessment does not measure the targeted skills taught in the class.

b) When three teachers score the same answer sheet, they assign different grades for the same answer.

c) When the context of assessment tasks is far removed from language as used in real life situations.

d) When the tasks used for assessment do not promote learning.

4. Which of the following factors is most likely to improve the reliability of an assessment?

a) Each answer sheet is scored by 2/3 examiners who use the same scoring key and set of scoring criteria.

b) The number of subjective-type questions is increased and that of objective-type questions is decreased.

c) All the assessment tasks carry equal weighting.

d) Each task should measure more than one major skill.

5. For assessing speaking skills, a teacher divides a class of forty into four groups. The first two groups are assessed during a day in two sessions: morning and post-lunch. The second two groups are assessed in the same manner, but on another day. The groups who were allotted the post-lunch session were unhappy with the teacher's decision because they used to get tired by then and they thought, therefore, their performance could get affected. So they wanted to be assessed during a morning session. But the teacher rejected their demand.

In the above situation, which principle of assessment is violated?

a) Validity   b) Reliability   c) Authenticity   d) Practicality

6. i) Convert the following sentences into Passive and rewrite them.

- I gave her a book.
- She drinks milk every day.

ii) Use the verbs (given in brackets) in the correct form to fill in the blanks in the following sentences.

- Originally, this novel (write) _____ in Hindi, but it (translate) _____ into Odia in 1985.
- A tree is lying across the road. It (blow down) _____ in the storm.

The above questions were used by a teacher to assess students' ability to use passive voice. The questions under 'ii)' are more than those under 'i)' in terms of:

a) Reliability   b) Practicality   c) Authenticity   d) Dependability

7. If assessment tasks are meaningful, interesting, relevant and engaging, it can be said that they are to a great extent:

a) Valid   b) Reliable   c) Authentic   d) Practical

8. A teacher wanted to design an assessment which would help her students diagnose their strengths and problems in relation to learning so that they could improve their language ability.

In the above situation, the teacher's concern is related to:

a) Validity   b) Reliability   c) Authenticity   d) Washback

9. A teacher designs an assessment in such a way that she gets a clear picture about the changes she needs to make in her approach to teaching to make it more effective.

The principle of assessment related to the above situation is:

a) Validity   b) Reliability   c) Authenticity   d) Washback

4

10. A teacher wanted to assess the speaking ability of his students. He divided the test into three sections. In the first section, he planned to ask each student to participate in a one-to-one short interaction with him; in the next section, they had to be engaged in a one-to-one long interaction with him; and in the last section, each student had to talk about a topic for two minutes. However, soon he realized that it might take too much time to complete the assessment.

The problem with the assessment described above is related to:

a) Validity   b) Reliability   c) Authenticity   d) Practicality

11. A teacher taught skimming and scanning to a class. Which of the following will be the most valid assessment of these skills?

a) A set of multiple-choice type questions based on a variety of texts which have not been used by the teacher in the classroom.

b) A set of *wh*-questions based on a single passage (not used by the teacher in the classroom) requiring students to answer each in 50 words.

c) A set of multiple-choice type questions based on a single passage which has been used by the teacher in the classroom.

d) A set of *wh*-questions based on a variety of texts (which have not been used by the teacher in the classroom) requiring students to answer each in 50 words.

12. A teacher teaches a class of students the pronunciation of the sound "/ə/" (schwa). Which of the following will be the most valid measure of the students' ability to pronounce the sound correctly?

a) The students are asked to interact with the teacher in pairs and the teacher makes them use the sound during the conversation.

b) The students' use of the sound is observed informally when they interact in the classroom.

c) The students' performance in pronouncing a set of discrete words containing the sound as well as their ability to use of the sound correctly in the classroom, both are taken into account.

d) The students' ability to use the sound while interacting in pairs with the teacher as well as the informal observation of their interaction in the classroom, both are taken into account.

13. A teacher taught writing leave applications to the headmaster to his students in a class. However, he assessed them by asking them to write a letter to either the District Collector or the Block Development Officer.

Which of the following principles of assessment gets violated in the above situation?

a) Validity   b) Reliability   c) Authenticity   d) Practicality

14. A teacher gave her students a few tasks to work on as a part of classroom assessment. Most of the students complained that the assessment was too difficult and that they were not familiar with the task-types used in the assessment.

Which of the following principles of assessment gets violated in the above situation?

a) Validity   b) Reliability   c) Authenticity   d) Practicality

15. All the students in a class fail to answer a question asked during an assessment. This implies that:

a) Students were not prepared well for the assessment and they need to prepare better for future assessments.

b) That particular question should not have been included in the assessment and therefore, all the students should be given marks for the question.

c) Such tough questions are meant to assess students' real ability to use language and thus, should continue to be included in all such assessments.

d) The whole class needs to be taught the item asked in the form of that question.

16. Which of the following practices of an English teacher at secondary level can be termed as unethical?

a) Letting students know about the skills to be assessed and scoring criteria.

b) Providing them adequate practice in using the skills to be assessed before the test.

c) Discussing the structure of the assessment with students before finalizing it.

d) Giving them in advance two of the most difficult questions to be included in the assessment.

17. Which of the following grading practices is the least effective in truly showing students' achievement?

a) The teacher wants all his students to submit their homework but he only grades the last one of each student.

b) The teacher uses students' performance in classroom quizzes and two major examinations to assign them final grades.

c) The teacher allows his students to work on their assignments several times if they wish to improve their grades.

d) The teacher takes into account students' classroom behaviour in addition to their academic performance while assigning them final grades.

18. A teacher does not know how to teach and assess listening and speaking skills and thus, the assessments given by the teacher in the class never include listening and speaking. Which of the following statements is true about the teacher's practice?
a) The students who are not so good at reading and writing but very good at listening and speaking are unfairly graded.
b) There is nothing unfair about this as the assessments include only skills that are taught in the class.

c) This kind of situation is common across the country and therefore, the question of fairness does not arise.

d) The teacher cannot be blamed for his lack of awareness about assessment of listening and speaking skills and thus, the practice cannot be called unfair.

19. Which of the following sequencing of steps is the best for designing classroom assessments?

a) Deciding the objective/s of assessment→taking a look at available resources→preparing sound assessment tasks→designing the assessment criteria→writing the scoring key→scoring and grading students' performance→giving feedback to students→declaring assessment results

b) Deciding the objective/s of assessment→designing the assessment criteria→taking a look at available resources→preparing sound assessment tasks→writing the scoring key→scoring and grading students' performance→declaring assessment results→giving feedback to students

c) Taking a look at available resources→deciding the objective/s of assessment→designing the assessment criteria→preparing sound assessment tasks→writing the scoring key→scoring and grading students' performance→declaring assessment results→giving feedback to students

d) Taking a look at available resources→deciding the objective/s of assessment→preparing sound assessment tasks→designing the assessment criteria→writing the scoring key→scoring and grading students' performance→declaring assessment results→giving feedback to students

APPENDIX C

<center>CLAL Questionnaire: Final Draft</center>

<center>Classroom Assessment Literacy Survey</center>

Dear teacher,

I thank you for agreeing to participate in this survey. The aim of this survey is to measure the classroom assessment literacy of secondary school English teachers in Andhra Pradesh.

Read the questions carefully before you choose your answer from the choices given. Please contact me immediately at 09160796502 (mobile phone) or <santosheflu@gmail.com> if you have any doubt about or difficulty in understanding any of the following questions. I assure you that your responses will be used only for research purposes and your name and other details will not be disclosed. Thanks a lot!

Best regards

Santosh Mahapatra, Ph. D. (English) Student, University of Hyderabad

Mobile No. 9160796502

Your name:

Name of the school and place where you work:

Age:

i) 20-30   ii) 30-40   iii) 40-50   iv) More than 50 years

Sex:

Teaching experience:

i) Less than 5 years   ii) 5-10 years

iii)10-15 years   iv) More than 15 years

1. The following assessment task (based on an unseen passage) was used by a teacher to assess students' (class VI to VIII) reading comprehension skills in a classroom. Look at it carefully and try to answer the questions given below it.

Read the passage and answer the questions that follow using full and correct sentences.

It's not just a fort, but a tiny little kingdom in itself, untouched by the ravages of time. Of course, the only alien giveaways are the plastic covers and litter strewn around everywhere. But learn to ignore them like you would do anywhere else in the city, and a visit to Golconda is like discovering a time machine.

The stories surrounding Golconda are many. If you like Bollywood masala in your ancient tales, your best bet are the tourist guides who swarm the entrance. They are full of stories that are intriguing, magical and most probably fanciful and true. But they are stories that will surely keep you entertained throughout your long climb up the fort.

Here's the true story. Golconda or "Golla Konda" (sheperd's hill) is a 13th century Fort, built by the Hindu Kakatiya kings. According to a legend, a shepherd boy came across an idol on the hill. This led to the construction of a mud fort by the then Kakatiya dynasty ruler of the kingdom around the site. In the 16th century, Golkonda was the capital and fortress city of the Qutub Shahi kingdom, near Hyderabad. The city was home to one of the most powerful Muslim sultanates in the region and was the centre of a flourishing diamond trade.

a. Why is Golconda considered to be "untouched by the ravages of time"? (30 words, 2 marks)
b. Why does the author ask us to ignore plastic covers and litter? (20 words, 2 marks)
c. What does the author mention about tourist guides? (20 words, 2 marks)
d. When was Golconda built? (10 words, 1 mark)
e. When did Qutub Shahis make Golconda its capital? (10 words, 1 mark)
f. Give a suitable title for the passage. (6 words, 2 marks)

---

i) Which skill/s does the task assess?

a) only reading   b) only writing   c) both reading and writing   d) only vocabulary

ii) If students do well on this test (7 marks or more), what will be the best inference?

a) They can read and understand the given passage well.
b) They can write very short answers appropriately and correctly as demanded by the question.
c) They can read and understand the given passage and write very short answers well.
d) They can read and understand any passage of similar difficulty level and write short answers well.

iii) Indicate whether the following problems are true (T) or false (F) for the above task.

a) The task does not assess what it aims to assess.
b) The task uses an unfamiliar context.
c) The task is too lengthy for students.
d) There is a strong possibility of having teacher bias in scoring the answer sheets.
e) There is little variety in questions (in terms of form).
f) Vocabulary-based questions are almost absent.
g) It may not be possible to administer this task during a classroom assessment.

iv) Indicate which of the following suggestions can (✓) and which cannot (×) make the reading comprehension task more accurate in measuring reading ability of students.

a) Giving MCQs to minimize writing.
b) Keeping only factual questions.
c) Adding a few questions on key vocabulary.
d) Adding some questions on the use of simple past tense.
e) Providing a scoring key with the task.

2. The following task was used by Ms. Ruth to assess students' (classes VI to VIII) ability to write paragraphs in a classroom test. Please go through the task and try to answer the questions that follow.

Write a paragraph on one of the following topics in about 100 words.
1. Your family   2. Your best friend   3. The person you like most
Total marks: 10 (Content = 3 marks, Organization = 3 marks, Vocabulary = 2 marks, Grammar = 1 mark, Overall = 1 mark)
Maximum time allowed: 30 minutes

---

i) State whether the following statements are true (T) or false (F) for the above task.

a) Scoring students' writing in this task involves very little possibility of teacher-bias.
b) Using the scoring criteria may lead to bias in the scoring of students' writing.
c) Students' low scores in individual areas like 'Content', 'Organization', etc. indicate their problem in those areas.

3. A teacher used the following task in a class test to assess students' (Class VII/ VIII) ability to interact in familiar situations. Look at the task carefully and answer the questions given below the task.

3

Given below is a dialogue, with a few responses missing. Fill in the missing responses:

Revathi: I'm really scared. I hope the teacher does not ask me to speak on my first day in this school.

Rashid: You don't really have to. The teacher...

Revathi: Tell me about your...

Rashid: Well, I still remember my first day here. When...

Revathi: I also wish to have...

Rashid: I'm sure you..

Revathi: Thank you!

(2 X 5 = 10 marks)

i) Which of the following skills does the above task assess?

a) only speaking   b) both speaking and writing   c) only writing   d) both writing and reading

ii) Which one of the following can be the most effective task for assessing students' ability to interact on familiar topics?

a) The teacher asks each student questions on a familiar topic.
b) Students interact in pairs on one of the topics provided by the teacher.
c) Each student speaks for one minute on a previously given topic.
d) Students are asked to write dialogues for an interaction between two persons.

iii) Which of the following seems to be the most effective way for the English teacher to get a thorough idea about students' progress and problems related to the learning of spoken interaction skills?

a) ask students to write an end-of-term test and record of their performance in that test.
b) collect information and maintain a record about both their classroom interaction with peers and the teacher, and performance in unit and term-end tests.
c) ask every student to maintain a record about their progress and problems and submit it to the teacher at the end of the term.
d) request all the subject teachers to grade students' ability to interact, and then assign an average grade to each student.

4. The following task was used by a teacher to assess students' (VI to VIII) ability to use subject-verb agreement correctly.

Fill in the blanks in the following sentences using correct form of verbs given in the brackets.

a. I _____ (decide) not to visit the museum.
b. All of them _____ (come) here to meet me.
c. Do you know when they_____ (come) here?
d. I _____ (decide) to quit my job.
e. Simi _____ (read) today's newspaper.
(1 X 5 = 5 marks)

---

i) Indicate whether the following problems are true (T) or false (F) for the above task.

a. The task uses a very limited context.
b. The task uses real life contexts in which language is used.
c. Having multiple answers to each question promotes learning among students.
d. Having multiple answers to each question creates confusion for students and scorers.
e. The task does not help in assessing students' ability to use subject-verb agreement correctly.
f. The task provides very good information about students' ability to use subject-verb agreement correctly.

5. Ms. Sameera assesses paragraph-writing skills of her class VIII students. She wants to give feedback to the whole class (40 students) about their performance so that they can improve their paragraph-writing skills. Which of the following ways can be the most effective one and at the same time, will take least amount of time? (Choose the best option out of 'a', 'b', 'c' and 'd'.)

a) She creates a grid in which marks are assigned for categories like CONTENT, ORGANIZATION, VOCABULARY, GRAMMAR, etc. and each category is further divided into something like the following:	
Organization (3 marks)	3= Can organize the main idea and the supporting details so well that they make a meaningful paragraph that is nice to read. 2= Can organize the main idea and the supporting details manageably well with only a few gaps in the paragraph. 1= Can organize some of the ideas with a lot of difficulty. 0= Needs to work hard to learn how to organize ideas in a paragraph.

b) She corrects each answer script, assigns marks to each student as shown below and hopes each student will learn from the corrections.

Roll no. of the student	1	2	3	4	5	6	7	8	9	10
Marks obtained out of total 10	7	8	5	2	6	4	3	7	5	4

c) She creates a grid in which marks are assigned for categories like CONTENT, ORGANIZATION, VOCABULARY, GRAMMAR, etc:

Roll no. of the student	Content (3)	Organization (3)	Vocabulary (2)	Grammar (2)
1	2	2	1.5	2
2	1.5	1	1	.5
3	1	1	1	.5

d) As it is a test, she assigns marks only for correct answers. She asks her students to prepare well so that they can score well in the next test.

Roll no. of the student →	1	2	3	4	5					
No. of right answers →										
No. of wrong answers →										

6. The sentences given below are some statements in relation to classroom assessment. Indicate which of them are true ("✓") and which are false ("×").

Ideally, classroom assessment should:

a) encourage students to learn the language skills.

b) help the teacher to identify the strengths and weaknesses of individual students.

6

c) help the student to know where they need to improve.

d) be part of the teaching process.

e) use real world contexts.

f) not involve students in the process of task selection, criteria of assessment and method of assessment.

g) not put students under pressure or anxiety.

h) take into account students' performance across an academic semester/ year while assigning the final grade.

i) make use of paper-pencil tests rather than getting into portfolios, journals, etc.

j) not include self- and peer-observation as modes of assessment.

7. Which of the following sequencing of steps is the best for designing classroom assessments?

a) Deciding the objective/s of assessment→taking a look at available resources→preparing sound assessment tasks→designing the assessment criteria→writing the scoring key→scoring and grading students' performance→giving feedback to students→declaring assessment results

b) Deciding the objective/s of assessment→designing the assessment criteria→taking a look at available resources→preparing sound assessment tasks→writing the scoring key→scoring and grading students' performance→declaring assessment results→giving feedback to students

c) Taking a look at available resources→deciding the objective/s of assessment→designing the assessment criteria→preparing sound assessment tasks→writing the scoring key→scoring and grading students' performance→declaring assessment results→giving feedback to students

d) Taking a look at available resources→deciding the objective/s of assessment→preparing sound assessment tasks→designing the assessment criteria→writing the scoring key→scoring and grading students' performance→declaring assessment results→giving feedback to students

APPENDIX D

CLAL Survey Scores

Serial Number of Teachers	CBSE	ICSE	State-Board	
			English Medium	Telugu Medium
1	14	16	13	22
2	16	17	19	16
3	22	12	15	17
4	17	15	24	20
5	18	14	13	15
6	19	18	15	24
7	13	25	26	26
8	20	22	23	18
9	33	31	31	30
10	24	18	18	20
11	19	17	28	19
12	25	24	26	27
13	19	14	25	23
14	23	22	29	25
15	18	13	25	22
16	31	28	15	20
17	17	26	16	28
18	15	16	28	14
19	23	15	13	14
20	24	14	27	12
21	28	25	30	29
22	24	24	19	27
23	27	25	16	13
24	32	32	33	14
25	26	25	27	19
26	18	13	14	16
27	31	24	14	29
28	20	17	12	28
29	29	25	17	14
30	25	21	24	13
Mean Score→	22.33	20.26	20.81	
Standard Deviation→	5.57	5.66	6.03	
Overall Mean Score→ 21.05 Overall Standard Deviation→5.86				

APPENDIX E

Teacher Beliefs Questionnaire

Teacher Belief Questionnaire

Dear teacher,

I thank you for agreeing to respond to this questionnaire. This questionnaire intends to capture your beliefs about classroom assessment of language ability. Kindly read the statements carefully before indicating your response. Thank you.

Best regards

Santosh Kumar Mahapatra

Ph. D. Student, University of Hyderabad

Name:

Age: Qualification:

Experience as an English teacher:

1. Have you had any training in language assessment during your B. Ed.?

a) Yes [ ]   b) No [ ]

2. How helpful and effective was the training?

a) Very helpful [ ]   b) Helpful [ ]   c) Not so helpful [ ]   d) Not at all helpful [ ]

3. Have you had any in-service training in language assessment?

a) Yes [ ]   b) No [ ]

4. How useful was it?

a) Very useful [ ]   b) Useful [ ]   c) Not so useful [ ]   d) Not at all useful [ ]

Please state your opinion about the following statements by selecting (√) one of the following options: SA (strongly agree), A (agree), D (disagree), SD (strongly disagree) and DN (don't know).

5. It is necessary to have training in language assessment to become an effective teacher.	SA	A	D	SD	DN

6. The first step in language assessment is understanding language proficiency.	SA	A	D	SD	DN

1

7. The test score should show the real ability of a student to use language.

SA	A	D	SD	DN

8. An assessment should not assess students' ability to memorize and reproduce information from their textbook.

SA	A	D	SD	DN

9. It is better to have regular classroom assessments than one final examination.

SA	A	D	SD	DN

10. Teachers should use a variety of texts outside the textbook to assess reading comprehension skills.

SA	A	D	SD	DN

11. Only those language skills which are taught should be assessed.

SA	A	D	SD	DN

12. While assessing language skills, learners should be informed in advance about the criteria on the basis of which they will be assessed.

SA	A	D	SD	DN

13. It is the duty of the teacher to inform students about why they get a particular grade/score for their performance in an assessment.

SA	A	D	SD	DN

14. Listening and speaking can be assessed by the teacher in the classroom even without a tape-recorder and an audio player.

SA	A	D	SD	DN

15. The teacher should not include such questions in the assessment that cannot be answered by him/herself without any immediate outside help.

SA	A	D	SD	DN

16. Enough context should be provided while assessing grammar and vocabulary.

SA	A	D	SD	DN

17. Those students who are slow at writing should be allowed a little bit of extra time if the assessment involves writing.

SA	A	D	SD	DN

18. Students should be given more than one chance to show their ability to use language.

SA	A	D	SD	DN

19. A classroom assessment is a great asset to teaching for a teacher.

SA	A	D	SD	DN

20. A good assessment encourages students to learn.

SA	A	D	SD	DN

21. A teacher should provide individual feedback to all students (even if they belong to a single large class) after an assessment so that they know their strengths and weaknesses.

SA	A	D	SD	DN

22. The teacher must understand the objectives of individual lessons and tasks to be able to construct effective assessments.

SA	A	D	SD	DN

23. The teacher should be aware of the basic principles of language assessment.

SA	A	D	SD	DN

24. The teacher should have the ability to analyse the results of the assessment and give proper report about the same to students, parents, school management, etc.

SA	A	D	SD	DN

25. The teacher should involve his/her students in the process of assessment.

SA	A	D	SD	DN

26. Assessment can happen while teaching; it may not necessarily be a separate activity.

SA	A	D	SD	DN

27. The teacher should consider factors like class, caste, gender and religion while designing assessment tasks so that it does not favour or give unfair advantage to any student/s.

SA	A	D	SD	DN

APPENDIX F

Rating Scale to Evaluate Task Quality

The following rating scale to evaluate the quality of classroom assessment tasks is designed specifically keeping in mind the objectives of classroom assessments proposed by the CCE. Four levels are created for each characteristic of the task. These levels are, in fact, points on a continuum with '1' representing the least fulfilment and '5', the best fulfilment of the characteristic or condition.

List of characteristics of the classroom assessment task	Rating Scale				
	1	2	3	4	5
1. It states the task objective/s clearly.					
2. It assesses what it intends to assess.					
3. It is/can be integrated with the classroom teaching.					
4. It provides/can provide information about students' ability to use the language skills and components covered.					
5. It gives/can give diagnostic information about students' learning.					
6. It requires students to respond to a simulated real life context by using their language ability.					
7. It motivates/can motivate students to learn.					
8. It leaves very little scope for teacher bias in scoring/grading students' performance.					
9. It offers/can offer every student equal chance to perform.					
10. It provides feedback to students about their performance.					

APPENDIX G

Pre-Intervention Teacher-Made Assessment Tasks

Teacher 1

Listening Skills.

clau IX .   15 min .

               To find specific information.

Task : Event Advertisements .

Type of questions : Fill in the blanks .

Marks : one mark for each question

out of 40 →   8   —   100%
                     20   -   70-80%
                     7   —   60-50%
                     5   —   20%

(margin note: Blank-filling / 5 types of question)

Feed back — 70-80% : extra careful
    with minute details .
60-50% — Need to be attentive -
5% — lack concentration

Speaking Skills

pronunciation, f   15 min.
                                    ──────
                                     10 m

Reading loud = un familiar text.

Marking scheme:

Intonation 2   Pronunciation 5   Expression 2

Language function : 1

Feed back →.

9"½ — phonetically correct, rising &
        falling tone, uses proper language
        function & expression.

7-8.   phonetically correct, forgets
        expression sometime

5-6 — pronunciation is average, does
        not express well

below 8 — Has mother tongue inflection
            cannot express

Writing Skills.   ▮▮▮▮ 10m.

Essay. Organisation, Accuracy, vocabulary.

IX . 1. Empowerment of girl child .
    2. Bad effects of watching Television .
    Marking scheme : Organisation - proper
    introduction, development, sequence and
    Conclusion .

            plight & necessity to empower a girlchild
            what can be done to empower
            a girl child - govt, family etc .
        Implications .

Content - $3\frac{1}{2}$m , Organisation - $2\frac{1}{2}\overset{3}{}$m, Accuracy, 2m ,
            Vocabulary $1\frac{1}{2}$

    2. Ill effects of watching Television .

        Introduction .
        Harmful effects of watching television .
        Conclusion - warning .

        Marking scheme :
            Content - 4m   Organisation - 3m Accuracy-2
            Vocabulary - 1 m .

Feed Back : students who score $9\frac{1}{2}$, 10 .
        starts with quotation, statistics & organise
        well, used suitable vocabulary without
        any grammatical mistake,
    ② 6-7 - Organises well, with few spelling
              and grammatical mistake .

⑧ m̄5-4 makes considerable mistakes not very
well organised.

below - Poor sentence construction & make
mistakes in spelling & grammar.

feedback 9½ - Needs to keep reading.

6-7 - Needs to take careg
spelling & organise.
collect some general information

5-4 - needs to follow cosol ✓

below 3 - Remedial writing.

Task1 - Class 8 - Reading Comprehension.
specific and global information.

what kind of question?
what kind of question task?
How will you answ?
who will get the highest mark / lowest (average)
How will you give the feed back.

Task

A medicinal value of Tulsi   5m

B. Biography of any great personality 5m

Question

+ Factual   Two question to test their understanding.
                                                    2x1 = 2m
ite   2   Two inferential questions.   2x1 = 2m
    answ limit
    3   word   one question based on vocabulary   1x1 = 1m

1 - Two question based on life history

2. Two question on achievements 2x1 = 2 m
                                   2x1 = 2 m

3. One question based on vocabulary to
    describe the person in short - 1x1 = 1m

Asseserent : Each question will be
awarded one mark. Since it's a Reading
Comprehension the child will not be penalised
for spelling and grammar.

who will get

Highest : 1) Around 20% students get 100%
2) Around 40-50% students get 70-80%
3) Around 20% students get 50-60%
4) Around 10% students get 30-40%.

70-80% - Need to be careful while writing inferential questions.

50-60% should read the paragraph thoroughly & pre-read the questions.

10% - should practice reading and understand the gist of the paragraph

3 80% → $\underline{80 - 100}$ ⇒ Task is easy

grammar.

Present perfect tense
during daily interaction.

· 10 m ·

Task : Paragraph (Report on any event).
· Visit & celebration description .
Paragraph with fill options (present
perfect or past).
one mark alloted to each blank.
out of 40
 6-7- 10/10 ·
 20 — 7-8/10 -
 :5 - 5-6/10·
 8 — len than 3/10 ·

<u>Feedback</u> : 20 - ~~Able~~ _{should be} to identify
  & use the tense ·

  5 — Application in integrated
    exercises is required -

  8 — Need to clear concept `

Vocabulary

prefixes & suffixes.

(Conversation) - Communicative

1. Exercises based on antonyms, adverbs, adjectives etc.

Type of question.

Fill ops — antonyms (in, un, non)

adverbs — change from adjective adverb.

MCQ —

marking scheme — one mark for each question.

Feed back — _lowscores_. learn the right prefixes & suffixes.

Average : Need to differentate and be specific with the usage !

Teacher 2

→. Listening.

→. Class – VIII

→ 15 min.

→ specific skills like – ability to find specific information.

→ Topic.

→ Speech By Abdul Kalam . – 15 min.

— Marks – 10.

→ Marks .

Questions
marks
Feedback.

→ Note taking . — quantity — 3.

→ Content. — 2

→ Topic – covered. — 2

→. Name the topic. — 1

→ what is all about / grammar – 1, Vocabulary. – 2.

→. Topic. covered.

→ How much child grasped.

→. How. the child answer the questions

Speaking skills

Pronounciation

→ Assessment.

→ Task.

— Feedback.

+ Time — 40 min.

× Each child — 2 min.

× A line of printed matter. → Reading aloud .

→. clarity | ~~~~ ⫽ pitches. | loud | stressing.
    2·5   2·5   2·5   2·5

+ Feedback.

→. Spontanity.

→ Clarity.

→ Emotions.

Task – 1.

Class. – VIII

Reading Skills.

* Read for a specific and global information.
* The questions.
* Task.
* assessment
* Highest. — lowest — Average.
   ↓
   mean.
* feed back.
* How will you use the scores

10 — 40%
10 – 40–70%
10 – 70 – 100%

→ Questions ✓

→. Who is the speaker.?
→ what did you understand from the passage?
→. What is the meaning of — —?
→. what is all about?
→. Enquire about the tense &
→ Grammar.

Task. ✓

→ I. Collect information from others on climbing mountains.
→ II. Why it is very difficult to over come the problem in life than the climbing mountains discuss. in pairs.

→ Assessment

→ Task 1 – 5 min. — listen to others — note —. Expressing — language.
   10 mks.   5 mk.   2.5 mk.   2.5.
→ Task 2 — 10 Min — How they get information / Answering Q. / lange. / long. flow.
   10 mks   from others (thinking)   2.5.   2.5.   2.5.
                       2.5 mk.

→ · Highest.
                  7.

— lowest.   10.

⊥ Average   23.

→ Feedback!
─────────

* Observation
* Speaking
* Spantainity.
* Answering.

* Writing skill.
  → how to write essay.
      → organization of ideas.
      → grammar.
      → Vocabulary.
*. 10 mks.

*

→ Write an essay on Hazards of Pollution. in 150 words.
                                                    — 8 mks

  → Introduction to the Topic. 1½ mks.
  → Definition. – 1.
  →. How / Why / usage of language } – 1. mk.
  →. Child's. resolution} 2½ mks
  → Conclusion. – 2.
  → Grammar. – 2

  →. Feedback.
  →. Hand writing.
  →. Structure.
  →. language.
  → usage of Vocabulary.
  →. How the child ended the task.

* <u>Grammar</u>.

→ Ability to use present perfect Tense. in daily interactions

→ 10 mks.

→. 20 min. time.

→ Printed matter. (a paragraph.) mostly in a regular. conversation in the class./ at home ~ ~ ~

→. 4 question on change the tense.

→ 4 questions M.C.Q.

→ 2. questions on. the fill in the blanks.

→. Feedback depends on how the child completes the task.

Vocabulary

→ Prefixes — suffixes.

→ 5 mks

→ A short passage - about 150 words.

→ give. some questions. based on vocabulary.

→ find out the meaning of the words from given passage

few from para I., II, III.

→. Read the following passage write the meaning of of the words given from the para I II III.

→ Assessment based on the child's performances

Teacher 3

Listening :–

Objectives :– Assessing Listening Skills.

<u>Assessment</u> <u>Criteria</u> :

8/10 – Best listener
6/8 – Good listener
4/6 – average
0/4 – below average

<u>Tasks</u> : Read out a story.

① put Question
② Ask them to tell the ghist
③ Note down important points
④ pulling the date/ events in chronological order

T.

Speaking ability.

Skills tested — ability to pronunciation
time : 40mts
class : IX

Give any article/reader's column, story book, ma
zine to the students ask to read out.
Observe each student when he reads the paragraph—
1) note down his pronunciation —
2) his ability to read, ✓
3) pause — 1 ✓
5) phonetics — 2 ✓
6) Stress — 1 ✓
7) intonation 1 ✓
8) Slang ! ✓
9) Pronunciation. 4
~~~~ Keep aside those who are substandard —
grade them — Give ranking—
Help those who are not able to pronunce by giving
Practice — Sugest to buy some C.D. Watch
Programmes in English in T.V. —. differentiate
between British Slang / American slang.—

Class IX:

1. Reading Comprehension:

2. Objective: overall Ghist - Specific information.
3. Time: 30 mts.   No of students: 30.

Plan of action:

① Give 2 paragraphs from any text book.
② Ask the students to read it ~~Silently~~ one by one.
③ Ask the students to Summarise it in few sentences
④ Ask Give at least 5 sentences to gather the Specific information of the two paragraphs -
⑤ ~~Correct the Scripts~~ Students answer orally -
⑥ Give marks based on their answers -
⑦ while reading note their Pronunciation Skills in reading -

How to plan for essay writing — How do you evaluate.

Steps:

(1) Choose any Current topic to the Standard of the children

(2) Allot 40 mts time ask them to write.

(3) Assessment.

Verify whether introduction,
Content —
Conclusion are there — or not

allot marks for each sub-heading — testing areas.

① legibility of handwriting —

② Content matter —

③ Grammatical ability —

④ vocabulary —

⑤ whether Current affairs are introduced or not

⑥ time - limit —

⑦ take out unfinished essays.

⑧ Valuate the Scripts —

based on the above give grades

Finally tell the students their mistakes —

_____

Introduction   Content —   Conclusion   Gramm.   Vocabulary
     1   5   1   2   2
                    3

                                                       legibility
                                                          ①

Vocabulary:—

Objective:— assessing Parts of Speech.

Assessment	Criteria
8/10	The child who is able to write all the forms of the given word.
6/8	who misses one or two
4/6	who does so). correctly.
0/4	Struggles to do everything.

task :— Give the pupil one word as hint to write down the different

Feedback : After the assignment correct answer will be written on the board & the child correct his answers — (Self scoring).

Grammar

Objective:- Assessing Tense forms.

Criteria:-

8/10 — The child who makes very few
mistakes is said to be the best ?

6/8 — A child who is able to answer 60-80 —
correctly

4/6 — able to answer only a few questions
correctly.

0-4 — not able to answer.

Task:- Fill up the blanks with correct form
of verb —

Feedback:- teacher writes the correct answers
asks the pupil to correct (self
scoring)

Teacher 4

Listening for specific information

1. I explain the story or paragraph to the Children.
2. Asked them listen carefully. I ask the question after it.
3. After explaining the topic.
4. I conduct one mark objective test.
5. I already prepare one question. I distribute to the children.
6. ~~Inst~~ Like that I can assess their listening skills.

Class - IX   Duration 40 Minuts.
            Asses pronunnevation to the children

1. Ask the children to read the text-book loudly one by one.
2. Observing carefully their reading skills
3. Through reading skills how he is pronouncing words also.
4. Phonetic sounds keenly observed.
5. Stress of the words.
6. Correct pronunciation
7. Silent ~~sout~~ letters how he is going to be reading observed.
8. Reading Constantly. Following he is stop at the full stop. also observed.
9. I prepare the ~~Statis~~ Statistical table writing the above points and ~~of~~ alloting marks.

Class - _IX_

Reading Comprehension : 1. Reading 2. Overall get and specific information

Students - 30

1. I would like to give the news paper to the children to read. random manner. If anybody reads fast giving the chance to another like that manner I asses the children to read the news paper. I give only three to four sentences.

In that manner everybody get opportunity to read. I assy reading skills for the children. Some of them may be slow readers eventhough I don't give much time to them but I identify to them they are slow reading children. After assessment I give one more chance to the readys who read fast. This time I give difficult paragraph to them in that manner finally I decide to assess the reading skills.

1. First of all I said to write the essay. which way given by me.

2. Concept may understand by the children or not I observe

3. The main important points are covered by them I observe.

4. First I ask them to write the important point after that to expand them I ~~let~~ told to the children.

4. I observe the sentence formation and Grammar.

5. They must imagine about the topic

6. important points they should note.

7. Keep all the points in Chronological order

8. Finally I asked them to expand the points.

9. They can discover the things what they imagined.

10. ~~the~~ In that ~~to~~ manner I assess their Essay writing skills.

Vobabulary collocation

1. Giving some collocation to the children asking them to write their own sentences.

2. ~~Giving~~ Asking them in so and so lesson identify and write collocations.

3. Through Structures we can assy the children.

Grammar   Subject- Verb   agreement-
40 minutes time

1. we prepare some examples for students
2. giving fill in the blanks or ~~and~~ choose the correct ~~th~~
    word.
3. I ask the children in this way.

Teacher 5

Listening

Objective: Assesing listening skills

Assesment criteria

Correct Answers   Wrong Answers

      2   - 0 -

Marks deducted for wrong spelling / grammatical
Task   errors.

(1) Students have to answer 10 questions

(2) The teacher reads them a story / news item.

(3) The teacher repeats if necessary.

Feedback — The teacher tells students orally about how
         They can improve their listening skills.

     Passage < more
               < less } difficult

Speaking

Objective — Asses speaking skills.

Assesment criteria

content	organization	pronounciation	Grammar
3	3	2	2

Task

students are given a topic in advance

They prepare and speak for two minutes

Feedback

Written comments and oral feedback
immediately after the student speaks.

Those who can't speak anything are encouraged

Reading

Objective — testing reading comprehension

Assesment criteria — correct answers — 2

spelling mistakes / grammatical errors
— -2

Task — Passage from the text book / unseen passage

Questions — (1) 'Wh' - questions

(2) Direct and inferential questions

Feedback

classroom revision

Writing

Objective – testing essay writing

Assessment criteria

content   organization   vocabulary   Over all
                            and   (spelling/style)

Task

Topics —— Population explosion in India

How you spent your last summer
                              vacation

Celebration of annual Dy in your school

Feedback

correction of grammatical, vocabulary and spelling
errors

comments (written) for improvement

Vocabulary

Objective — to asses how much students know about using antonyms and synonyms.

Assesment criteria

	Right use	Wrong use
Words	- 1 -	- 0 -
Sentences	- 2 -	- 0 -

Task

total 15 questions

10 words

5 sentences

Feedback

Questions are discussed in the class room after the test.

Tips are given to the students

Grammar

objective — to asses how students use articles

Assesment criteria

Right use	Wrong use
-1-	-0-

Task — total 10 questions
- blank filling
- 10 sentences

Feedback — correct answers are given to students after the test.
The use of articles is discussed in the class room.

Teacher 6

LISTENING

Objective :- Assessing Listening

Assessment criteria:-

1. Correct answer - 1 mark
2. Wrong answer - 0
3. Spelling errors - 1 mark minus for 3 errors.
4. Summary

    * Excellent   * good   * Average   * Poor
       100%   up to 80%   up to 60%   below 40%

Task :- Based on a story or a news item read out by the teacher

    1. Listen and fill in the blanks
    2. Listen and ~~summarise~~ summarise

Feedback :- Oral feedback through discussion

SPEAKING

Objective :- Assessing Speaking

Assessment Criteria :-

1. Excellent Speaking — No grammatical error,
   No Prounciation error,
   Good organization.

2. Good Speaker — Very few grammatical and
   pronunciation problems.
   Good Organization.

3. Average Speaker — Regular problems with
   grammar, prownciation and
   organization

4. Poor Speaker — Does not know how to speak

Task :- Group discussion/Speaking on a topic

1. Decision by the teacher how ~~much~~ many marks
   to be given to each student

Feedback :- The teacher tells the class about how to improve
   speaking

READING

Objectives :- Assessing Reading

Assessment criteria :-

1. Reading a passage and answering all the questions based on the passage correctly — 80% to 100%
2. Only 60% of the questions answered correctly
3. 40% of the questions answered correctly
4. Doesnot know how to read.

Task :- Passage from the text book or news article or such other sources.

1. Mixture of Wh - questions, inferential questions, questions on global information.

Feedback :- Correction by teacher and oral discussion

WRITING

Objective :- Assessing letter/~~Rep~~ report / article / essay writing

Assessment criteria :-

1. Can write without any spelling, grammatical, vocabulary and format errors
2. Can write but with some errors
3. Writes with difficulty
4. Can not write anything properly

Task :- The teacher gives three topics to write on. These topics can be for a letter, report, article or an essay

Feedback :- Thorough checking and corrections of every script

VOCABULARY

Objective :- Assessing Vocabulary — Word meaning, antonym/synonym, collocation, prefix/suffix

Assessment criteria :-
1. Included with reading assessment
2. Oral assessment in the class
3. Right answer — wrong answer — half-correct answer

Task :- Using words and in sentences, matching exercises, blank-filling, etc..

Feedback :- Correction and classroom discussion

GRAMMAR

Objective :- Assessing Grammar - Direct/Indirect speech, Active/Passive voice, Tense forms, etc...

Assessment criteria :-
1. At Right - Wrong answers
   1 ——— 0

Task :- Error correction, blank-filling, sentence completion, sentence transformation, etc...

Feedback :- Correction and classroom discussion

APPENDIX H

Post- Intervention Teacher-Made Assessment Tasks

Teacher 1

Listening.

IX   - 15 min

Find. specific information from oral text.

Materials : An interview with APJ Abdul Kalam.

Level   : Highest: Gets needed information.
He/she is able to gets specific information he/she is required to find.

Lowest : He/she needs to work hard to improve his/her ability to find specific information.

Average: He/she is able to get the required information to some extent and has some difficulty

Assessment Criteria.

Grade A   8-10   - Almost no mistakes.
Grade B   - 6-8   - few mistakes.
Grade C   - 4-6   - Some mistakes
Grade D   - 2-4   - Lot of mistakes
Grade E   - less than 2 - Unable to get information.

Task:

1. Blank filling.   a. worksheet
2. MCQs.   b. Play the audio
                    If necessary play twice

Marking scheme : one mark for 1 blank / question.

Feedback :

use levels & assessment criteria and a remark regarding area of improvement.

Speaking

Aim: To assess ability to pronounce war

Class: X

Level description: w

· Level 1 : He/she pronounces/uses correct pronunciation for single words and sentences.

Level 2 : He/she pronounces single words and uses correct pronunciation in sentences with very few problems,

Level 3 : He/she pronounces single words and uses correct pronunciation in sentences with some problems.

Level 4: He/she faces difficulty in pronouncing words regularly.

Level 5: Has to struggle a lot to pronounce single words

Assessment : 8-10 - He/she is able to pronounce all the words & sentences correctly

6-8 - He/she is able to pronounce most of the words with some difficulties.

2-0 : He/she has problem in pronouncing even the single word.

Task :

1) Group Discussion / JAM (In class)
2) List of Role play (In class) .

Scoring Key

(10m)

6m - correct pronunciation in single words
4m - correct pronunciation in sentences

Feed Back (oral) .

Suggestion in Improvement in pronunciation .

Reading Skill

Class X

Aim: 1) To find specific information.
2) To find gist.

Level Description

Level 1 : when He/she reads without any
problem and get the gist of the text
and specific information

Level 2 : Can get the specific information
and gist but has few problem.

Level 3 : Can do all there with quite a
few problems.

Level 4 : He/ they needs to improve
reading ability.

Criteria :

8-10 : with almost no problem in reading

6-8 : with few problems

0-4 : Has problems in reading &
understanding the text.

Task

1) Advertisement
2) Discursive Passage

1) MCQs
2) True/False, Y/N.

No 'wh' questions.

Scoring key

1 mark for each correct answer.

Feed Back

1) Oral
2) written

Through classroom discussion.

Writing

Duration : 20 min.

Class IX

Essay writing :-

Organisation, accuracy, Vocabulary

Topics :
Level I :
   A₁ : content is relevant, organises well & is coherent in writing. Writes correct sentences with appropriate punctuation & vocabulary.

Level II : He/she writes with a little problem in organisation of words opr statemind) problems in accuracy & vocabulary.

Level III : He/she faces regular problems in organising, struggles to corik correct sentences and gropes to find appropriate words while writing.

Level IIII : Needs to work very hard to organise ideas & present them ronter simple sentences & does

Score :   Organisation-4 , Accuracy-4   Vocabulary-2

A — 8-10
B — 6-8
C — 4-6
D — 0-4

Topics : ① Pollution in your locality . Measures
        to Combat it .

    ② vegetarianism is a way of healthy
       life . Agree / Disagree elucidate .

    ③ Social networking — Boon or a Bane .

Feedback : written and oral :
    → organisation underline the grammatical
    mistakes and encircle the spelling
    mistakes .

Result : calculate average scores . to find
    the lacunae .

Vocabulary

Duration: 10 - 20 min

Word approach:

Understanding the meaning and usage of the familiar words.

level I: she/he understands the meaning of the word in different context and use the word orally and writing.

level II: Has mind problems in using the words as per the context.

level III: who faces regular problems in understanding and usage of the words.

level IV: who has to work very hard to understand the meaning & usage of the words.

Scoring: Three marks for each word

1 for meaning
2 for sentence.

Feed Back: Elicit/ Classroom discussion

Oral feed back

Interpretation: Grill form in usage & ~~mistake~~

meanings.

Grammar.

Reported speech in Communicating

level I : uses reported speech while speaking & writing without any difficulty.

level II : uses reported speech with some minor difficulties.

level III : Regular difficulty in using reported speech in writing & speaking

level IV : needs to work really hard to learn how to use reported speech.

scoring : 10 m :

        Grade I - 8-10
           II - 6-8
           III 4-6
           IV 0-4

Using Reported speech
   5- speaking   5- writing.

Task: Conversation
   ① Writing - Passage
   ② Speaking - Report (orally)

Feedback:
   Tense form wise feedback.

Results: Focus on tense (speaking)
   writing :) sentence form.

Teacher 2

Listening.

Class - 8.   15 min.

* Ability to find specific information from an oral text

→ Materials - Speech by Sharukh Khan.

→ Level
  → Highest - nf. he/she is able to get/find the specific information he/she is asked to find.
  → Lowest - Somebody who needs to work hard to improve his/her ability to find specific information

→ Assessment criteria:

Grade→ 8 - 10 : Almost no mistakes.
  A

Grade - 6 - 8 : a few mistakes.
  B

Grade - 4 - 6 : Some mistakes.
  C

Grade Below 4. : Lot of problems./unable to get information
  D.

Task.

→ Multiple choice. questions.

→ Worksheet.

→ Scoring key.

→ Feedback — use levels., assessment criteria, a remark regarding area of improvement.

Speaking. — To assess ability to pronounce

Class - VIII

Time:

→ Level description.

Level 1:- Somebody uses correct pronounciation for single word and sentences.

Level 2: Somebody who pronounces single word and uses correct pronounciation in sentences with very few problems.

Level 3: Somebody who can do with very some problems.

Level - 4" Who faces regular difficult in pronouncing words.

Level - 5: struggle a lot to pronounce a single word.

Assessment Criteria

8-10 → He/she able to pronounce all the words & sentences with - out any difficulty.

6-8 → He/she able to pronounce words with some difficulty.

06 → He/she struggle to pronounce even a single word.

Task: 40 min.
→ group discussion / a list of roll plays. — ask them to choose. and speak in class.

Scoring Key.   Marks — 10.

→ 6. mks. — fr. Correct pronounciation of words.
→ 4 mks — for Correct pronounciation of sentences.

Feed Back.

— Mostly oral.
→ How they can improve.

Reading Skill.

→ To find specific information and gist.

→ Level description.

Level 1. → Who can read without any problem and obtains
specific information and gist of the text

Level - 2. Can get the specific information & gist with a
few problems

Level - 3) Can get the specific information & gist with quite
a few problems.

Level - 4 Needs to improve his/her. reading ability.

Criteria assessment.

8 - 10 - 6 - for specific with very few / no problems.

6 - 8 - few problems.

9 - 4. Lot of problems / struggles to read text.

Text : Advertisement / a story / article / a passage

Questions : MCQs / yes / No., True / False.

Scoring key : 1 mk for each correct answer.

Feed Back : Both oral / discussion (group).

* <u>Writing</u>

Time - 20 Min.

→ Essay writing skills. — Organisation.
- Accuracy.
Vocabulary

→ Level Description.

<u>Grade 1 :</u> Somebody who who shows coherance in writing and
8-10. organizes ideas well, writes correct sentences with
appropriate punctuation and vocabulary.

<u>Grade-2 :</u> Few problems in organization, some minor problems
6-8 with vocabulary.

<u>Grade - 3 :</u> Somebody who faces regular problems with organizing
4-6 ideas, writing correct sentences, and gropes gropes for
appropriate words while writing.

<u>Grade -4 :</u> Needs to work very hard, to organise ideas, to write
0-4. Simple sentences, and use a few familiar words.

Organisation — 4/10.
accuracy — 4/10.
Vocabulary — 2/10

<u>Topics :</u> * A. place recently visited.
* Cleanliness in locality.
* Use of mobile Boon/Bane.

Feed back ; written. +. Oral.

→. Arrows for organisational mistakes.
→ underline the gramatical mistakes.
Result : calculate average scores in organisation. *

→ Vocabulary. / 20 min - Time.

Class - VIII

→ ~~Word~~ understanding the meaning and usage of a word. (familiar)

→ Level Description:

Grade 1 — Somebody who understands the meanings of the word
8 - 10 — in different contexts and also use the word orally
and writing.

Grade 2 — understanding the meaning with little difficulty also
6 - 8 — usage in writing & speaking

Grade - 3 — ~~Faces~~ Faces the problem in understanding the
4 - 6 — meanings of a word.

Grade - 4 — works very hard in understanding the meanings
0 - 4 — of a word and use it

Scoring: — writing —

correct formation of sentence - 2 } 3 mks.
meaning - 1. }

feedback: Through classroom discussion.
oral feedback

Interpretation: calculate percentage.

VIII Class.
* Grammar.

* Active voice / Passive voice. in writing.

* Description:

Grade - 1. somebody who able to use voice in writing and
8 - 10   speaking without any difficultly

Grade - 2: somebody who able to use voice in writing and
6-8   Speaking with some minor difficulties.

Grade -3: somebody who faces regular difficulty while using
4-6   voice in speaking and writing.

Grade -4: somebody needs to work very hard to learn how
0-4.   to use voice in writing and speaking.

Scoring Key: writing
    using passive   writing - 5/10. speaking - 5/10

Task: writing.   Speaking:
    Procedure passage.   Report orally. Procedure orally.
    Blank filling activities.

feedback: Tense wise feedback.

Result: Focus on Tense. speaking / writing.
    Sentence problem.

Teacher 3

Listening / Speaking.

Objective:- Listening for Specific & global information. | Speaking Skills.

Assessment Criteria:-

Listening. (50%) — For Specific information - 50%.
For global information 50%.

Speaking (50%) Content — organisation - Pronunciation Vocabulary Grammar.

Task:- using the text-book.

involving Students in oral discussion- —

keeping individual record-

assigning grades-

Feedback. Individual | oral & written

Reading:

Objective:- To assess whether the students are able to read & understand a given passage for Comprehension.

assesment Criteria :-

8/10 — Students having no problem in reading & Comprehending it and get required information

6/8 — All the above with some problem

4/6 — Could get understands the text partially and get some information from the text.

0/4 — has to improve reading for Comprehension.

task :- passage to read / ① Any advertisement ② any speed by leader ③ an article on certain topic.

Questions — MCQ, true/false - arrange Scable sentence.

Feedback: oral discussion - Peer Correction.

Evaluation :- If the students fail in answering the given test ie if modification is necessary -.

Writing

obj - Assessing Essay writing Skills.

Assessment criteria :- If Continuous are

① Content :-	② organisation	③ grammar & vocabulary.
Formative one		
Formative two		
Formative three		

[Conduct 3 tests - across the academic year & evaluate the ability & progress of the child].

Feedback :- Individual feedback - written Scripts - Some times oral feedback. in mother tongue

Vocabulary: Integrated with reading

Task: New words / unfamiliar words from the text book

Tasks to be given:

① Guessing meaning.
② Antonyms & Synonyms in sentences
③ Mind Mapping
④ Substitution tables —

Grammar :-

Assessing any one grammatical component.

Criteria of Assessment — Can be done
Conducting writing Or Speaking - [integrated].

writing :- [From their text books and writing assignment take any one grammar point-

Speaking — Focus on one grammar point every week.

Scoring :- Assigns grades on the basis of correct usage of grammar. [Should be overall]

Feedback. Individual - [writing for best, & above average)

Slow learner :- oral feed back.

Teacher 4

Listening :- objectives should be

1. Assessing students ability to understand the the main idea.

Assessment criteria :- 1. observation

observation 1	observation 2	observation 3
A/B/c/D	A/B/c/D	A/B/c/D

Grade - A - Able to get the main idea. without any problem.

Grade - B - Able to get the main idea with some minor problems.

Grade - c - Able to get the main idea but faces lot of difficulty.

Grade - D - should work hard to listen and understand the text.

Feedback :- based on observation Grades. Students will note their problems and how to improve.

Speaking

Objectives – To assess speaking skills
(can be assess with listening)

Assessment Criteria

Voc/Grammar	Content	organisation	pronunciation
FA 1			
FA 2			
FA 3			

t. ~~vocabub.~~
g will write about students progresses in all three
categories. In all three assessments. Assigning
the grade on the basis of performance of the
students in the assessments.

Feedback :- Integrated with assessment reports

Task – student - student class room discussion .
                          or
student – teacher – interaction
                    or
student preparing and speaking about the
topic.

Reading And Vocabulary.

Objective - Assessing students skimming and scanning skills.,

Assessment Criteria:-

	Skimming	Scanning	Vocabulary
FA 1			
FA 2			
FA 3			

9 assessing their skimming and scanning and vocabulary skills in three assessments and writing a report. Grady are assigned on the basis of the progresses.

Feedback:- Both oral and written and includes suggession for improvement.

Writing

Objective :- Assessing formal letter skills.

Assessment criteria.

	Content 40%.	Organisation 40%.	Grammar and vocabulary 20%.
FA1			
FA2			
FA3			

Students performance is assessed and grading has done. Final Grade is assigned on the basis of progress.

~~Feed back:~~
Task :- ~~Individual~~ Individual / Pair / Group

model are given

students are asked to follow the models and write letters.

Feedback :- Feedback is given to each student after assessment. Mostly written sometimes oral.

Grammar

Objectives :- students ability to write and speak correctly.

Assessment Criteria

	Integrated with writing and speaking
FA1	
FA2	
FA3	

Feedback :- since the assessment is integrated I will focus on one Grammar item in speaking or writing.

Teacher 5

Listening

Objective : listening for particular information

Assessment criteria:

Level 1 — listens, understands and obtains particular information while listening:

1) without any difficulty
2) with some minor difficulty
3) with some amount of difficulty
4) Needs to improve listening skill

Task : story / news item / advertisement / song

1. Question type: blank filling / information transf
   multiple choice

Feedback : group discussion / student-student
   feedback

1. monthly / after formative assessment

Record : audio-record maintained for each
   student.

Evaluation: audio- change regularly, question-
   change after each assessment:

Speaking: Student's presentation skill

Objective

1) organizing content
2) pronounciation
3) Vocabulary and grammar
4) pace of speaking

Assessment criteria

Grade	Organization	Pronuncia-tion	Vocabulary grammar	Pace
Level A	Shows a lot of coherence	uses accurate pronounciation	uses correct sentences and words	maintains appropriate pace in speech
Level B	Some amount of coherence	minor problems with pronounciation	uses correct sentences and words on most occasions	on most occasions
Level C	struggle to organize	Pronounces some words correctly	on some occasions	on some occasion
Level D	needs to work hard on organization	needs to improve a lot	needs to improve in using correct sentences/vocab	needs improvement

Feedback: student-student / group discussion for weak students, individual feedback by teacher.

Task: topics from science / history / geography / literature

evaluation: Through discussion with students

Reading

objective: reading to know the gist of the passage
to know the organization of the text
to guess the meaning of the words from
the passage

Assesment criteria

Grades	Understanding of the gist	Text Organization.	Guessing word meaning.
Level A	gets the gist without any problem	finds the main idea, supporting points easily	guesses most unknown words
Level B	with some problem	with some minor difficulty	Some of the unknown words
Level C	struggle to get the gist	struggle to see main idea and supporting points	only a few unknown words
Level D	needs to work hard to read and understand the passage	needs to work to understand the organization of the text	needs to learn how to gues

Task: unseen story/newspaper article/advertisement/letter/poem

Questions: multiple choice/true/false/yes-no/information transfer

Feedback: classroom discussion

Teacher talks to weak students

Writing

Objective: Assessing formal letter writing skills

Assessment criteria

Levels	Content	Organization	Format	Vocabulary & Grammar
3 A	Includes all the necessary information	All the info and ideas presented in a very well organized way	appropriate format is followed at every stage	uses appropriate vocab and ~~current~~ using current
2 B	All necessary and irrelevant information - only a few things missing.	On most occasions organized.	with minor problems	on most occasions and with minor problems.
1 C	Some necessary information	Organized - only on a few occasions.	only in a few places	On some occasions and regular problems
½ D	needs to to include a lot of necessary information	needs to organize ideas and info	needs to follow appropriate format	needs to use appropriate words and correct sentences.

Task — write a letter to the counciller of your ward to take steps to control mosquito problems in your locality.

or

write a letter to the editor of a newspaper about a social issue that is affecting people around you.

Feedback—

I will ask students to keep a copy of their writing in a folder, which will be a collection of writing samples of students produced at different stages.

Grammar

Objective — Assessing how students use subject verb agreement

Assessment criteria

Levels

8 — 10 A - Almost always uses correct subject-verb agreement while writing.

6 — 8 B — On most occasions correct s/v agreement

3 — 6 C — On few occasions correct s/v agreement

0 — 3 D — Needs to learn s/v agreement

Task — (1)
   (a) correct the errors
       can be checked in the writing tests
Feedback — classroom discussion /written
       — exchanging answer scripts among students
evaluation — If too many students go wrong in this
       area, I will focus on it while teaching.

Teacher 6

LISTENING & SPEAKING   40 minutes
                        20 marks.

Objectives:- Assessing:-

1. Listening for particular information.

2. Listening for overall information.

3. Asking questions to get information to

4. Responding to questions.

Assessment criteria:-

		Listening 1	Listening 2	Speaking 1	Speaking 2
Excellent	A	Obtains particular information while listening without any problem	Can find the main idea of a listening text without any problem	can ask appropriate and correct questions to get information	can respon to questions appropriately
Good	B	with some minor problems	with some minor problems	with some minor problems	with some minor problems
Average	C	with some amount of difficulty and struggle	→ "	finds it difficult	with a lot of difficulty
Needs to improveD a lot		Needs to learn. how to listen for particular information.	How to "listen for main idea	Can not form questions.	can not respond to questions.

Task :- The teacher reads out a news item to the class

- The students respond to the questions posed by the teacher

- Then the students are divided into two groups. One group asks questions and the other group responds and vice-versa

Feedback: Both students and the teacher comment on the performance of students.

READING & VOCABULARY

Objectives:- Assessing :-
1. Skimming and Scanning skills
2. Infering word meanings from a reading passage

Assessment Criteria :-

	Skimming	Scanning	Word Meaning
Excellent A	Can read a passage and understand its overall meaning without any problem	Can read a passage and obtain specific information from it.	Can infer meanings of unknown words in a text
Good B	with some minor problems →	"	"
Average C	with some major problems →	"	→ "
Needs a lot of improvement D	Finds it difficult to read.	→ "	Can not infer meanings of unknown words from a reading passage.

Task:- An unseen passage of students' standard.
— Multiple-choice Questions
— Yes – No/ & True – False
— Jumbled Sentences

Feedback:- class room discussion after the assessment

WRITING

Objective :- Assessing essay writing skills

Assessment criteria :-

Assessment Criteria :-

	Content	Organization	Format	Grammar
Excellent A	Includes all the required details	Main idea and supporting details properly organized in the paragraphs	Very accurately followed	Writes all correct sentences
Good B	Most of the required details	Paragraphs are organized but coherence required	on most occassions	Most sentences are correct
Average C	Some of the required details	Needs to organize paragraphs well and maintain coherence	on some occassions	some sentences are correct
Needs to improve a lot D	Needs to include relevent and details.	Needs a lot of improvement in organization of ideas.	Needs to learn to follow appropriate format	Needs to write correct sentences

Task :- Essay writing
   1. Individual/pair
   2. 2/3 drafts
   3. Topics given by teacher - Independence day celebrations
   4.   School day
                                      A visit to a place

Feedback : Written feedback on one aspect of assessment
             criteria in one draft.

GRAMMAR.

Objective :- Assessing students ability to use reported speech

Assessment criteria :-

	Reported speech use
A	Can use reported speech effectively in speaking and writing
B	With some minor problems
C	With some regular difficulty
D	Needs to learn ~~what~~ how to use reported speech

Task:- Story telling from dialouge scripts/can be writing
1. Individual/pair activity
2. Students collect dialouge scripts.
3. Teacher monitors.
4. Exchange of scripts between pairs.

Feedback :- Student — student correction + teacher correction
Both oral and written feedback.

APPENDIX I

Sample of Teachers' Diary Entries

Teacher 1 **(Week 1)**

With lot of problem, I joined this training. So I wanted to use the opportunity for learning. Assessment is a crucial fact in our lives. During the first week the instructor developed my realization about many things. I know the meaning of language proficiency and content subjects. But I didn't have clear idea. We discussed these words and concepts. Fortunately we understood the meaning very well. I am expecting to use this knowledge for my teaching.

The word 'classroom assessment' was new to me. The fifth task showed one common problem teachers face. The question paper is a big problem. If the teacher knows how to prepare question paper this problem can be solved. This kind of training will empower us. All teachers must be given training.

Teacher 2 **(Week 2)**

I could not digest everything in my first week. This is new thing. I was not given this training. I am basically a science teacher. My colleague made me to understand everything in simple language. But I understood many concepts a little. But I felt more comfortable this week.

This week was fantastic. I got more support from the instructor. I learnt to set teaching goals. It was a realisation. I want to assess skills only those I have taught. I will divide these skills into parts and give marks to students for each part. Every teacher must know about sub-skills. Then only he can use assessment criteria. I will use it in my class if I get time.

Teacher 3 (Week 3)

I am always worried about practical problems. Assessment gives more head ache. In training programmes everyone talks about same old thing. We never discuss about practical problems teachers face. We have to manage ourselves only. This training programme gave us chance to discuss about the practical problems, options and principles of assessment.

I learnt many useful things. Resources, methods of assessment and principles of assessment are pivotal concepts. The tasks were not easy at all. The second, third and fourth tasks were tough. But I think I can acquire more knowledge if the tasks are tough. The trainer's presence made the difference. He guided us. I never lost my way and deviated. I clarified all my doubts then and there. I would not use my brains if I had easy tasks. I hope to continue well.

Teacher 4 (Week 4)

I found it very difficult last week. All tasks were difficult. Both of us had problems. I requested the instructor to make it easier. I want to gain knowledge why because I can do my duty better. So the he changed and this week was much better. I took a lot of time to prepare tasks but there were many new things for me.

I want to know more about CCE and how to make good tests. So I liked this week's tasks. They were more practical. I had discussed about using this information with my colleague. They were easy and interesting like distance education materials. So I did all the tasks myself. I hope the future weeks will be like this.

Teacher 5 (Week 5)

Very few teachers think about analyzing and interpreting assessment results and providing required feedback. These are still some of the luxuries a teacher can live with. I do not wish to be a part of that kind. So I took the areas focused during the last week very seriously. It was quite an experience after learning how to design tasks.

At the outset, I had just one query: How will it help me in improving my students' language skills? By the time I had completed the tasks, I realized that I could figure out quite a few things about my students' progress from their scores. Providing feedback is not any more difficult from here. I am almost certain that I can contribute to their progress directly through assessment.

Teacher 6 (Week 6)

I knew that it was the last week of training. It was quite relieving, but at the same time it also marked the end of a fruitful opportunity. I made the best use of the sessions asking questions and clarifying doubts. I made a list of questions and asked them to the instructor. I felt that alternative assessment methods and evaluation of assessments will be used by me quite soon.

The educational boards want us to use alternative assessments. But we need more discussions. On the one hand, it is important to learn about these methods, on the other hand, the practical constraints must be thought about thoroughly. I have little disagreement with the instructor that these methods will help students learn better. But the principles of the board and the school don't allow that freedom.

APPENDIX J

Syllabus for Intervention

Day	Syllabus Content
1	I. Construct of language proficiency II. Why assess language proficiency and not prescribed content III. Classroom/teacher-prepared assessment of English language ability
2	IV. Deciding objectives of assessment V. Developing assessment criteria
3	VI. Considering available resources VII. Choosing assessment methods VIII. Basic principles of assessment
4	IX. Developing tasks for assessing LSRW, Vocabulary and Grammar
5	X. Analysing and interpreting assessment results XI. Providing feedback to students
6	XII. Alternatives assessment methods XIII. Evaluating and improving used assessments
Total: 18 hours	

APPENDIX K

Tasks Used for the Intervention Programme (Sample)

Week 1

Task 1 Group Discussion **(30 minutes)**

Aim: This task aims to enable the participants to reflect on what they teach in their classroom and how far their teaching is geared to improving students' language proficiency.

1   The participants are asked to think about their own teaching in the light of the following questions:

- What are the objectives of your teaching in the English class?
- How does your teaching benefit your students?
- How do you expect them to use the language skills?
- Who is a successful language learner?
- How do you relate students' examination/test/assessment performance with their ability to use language in other situations?

2   All the participants are allowed some time to respond to these questions. Then they share their opinions individually with other participants and the programme instructor. Finally, the instructor winds up the session by emphasizing that they should focus on developing and assessing the language proficiency of their students in the classroom.

Task 2 Individual Work and Group Discussion **(30 minutes)**

Aim: This task tries to make the participants re-think about the concept of "language ability" in relation to their students. It also attempts to help them understand and define language ability.

1   The participants are asked to work in pair, look at the following diagram and give a description of four levels (1, 2, 3 and 4) of language ability and at least two of the corresponding levels in LSRW, Vocabulary and Grammar. They are given 15 minutes to get the descriptions ready.

```
┌──────────────────────────────────────────────────────────┐
│   Language ability: Level 1 (Advanced), Level 2 (Good),   │
│           Level 3 (Average), Level 4 (Below average)      │
└──────────────────────────────────────────────────────────┘
```

| Listening | Speaking | Reading | Writing | Vocabulary and Grammar |

2   Each pair presents their descriptions to other participants and the instructor. A group discussion follows after all the pairs complete their presentations. The instructor guides the discussion and ensures that the participants understand the concept of language ability by the end of the task.

Task 3 Individual Work and Group Discussion (30 minutes)

Aim: This task intends to raise awareness among the participants about the differences and similarities between assessing students' achievement in the English classroom, in terms of both skills and content, and their language proficiency.

1   The participants are asked to think about the last classroom assessment they have conducted and fill in the relevant details in the following table individually.

The purpose of the classroom assessment	
What did you assess?	
How did you assess?	

What information did you get?	
How is it different from tests like IELTS or TOEFL?	

2   The researcher examines the responses provided by the participants and raises leading questions to help them understand the differences between assessment of students' achievement in terms of language skills and textbook content. Effort is also made to enable them to see the differences and similarities between students' achievement and improvement in proficiency.

Task 4 Group Discussion (**30 minutes**)

Aim: An effort is made to enable the participants to realize the importance of assessing language ability in all classroom assessments.

1   The participants are asked to look at the diagram presented below, respond to it and comment on others' responses. The researcher monitors the discussion.

Classroom teaching: a story "Lion and the Mouse", a poem "Daffodils", an essay "Of Studies"

End-of-chapter reading comprehension, writing, listening, speaking, grammar and vocabulary exercises are practised in the classroom.

What to assess at the end of the lesson to know about the effectiveness of your teaching and their learning?

2   The researcher ensures that at the end of the discussion, the participants understand the importance of assessing language ability in all classroom assessments.

Task 5 Individual Work and Group Discussion (**30 minutes**)

Aim: This task provides a platform to teachers to understand the difference between teacher-prepared classroom assessments and those prepared by others and supplied to schools from outside. Their attention is also drawn towards teacher-readiness and teachers' ability to assess in this connection.

1   The participants are asked to go through the assessment situations presented below and point out what they think about each of them.

Situation 1: A group of students from a rural government high school (Telugu medium) in Anantapur takes its mid-term examination. In the English paper which was supplied by district officials, they are asked to write an essay on a set of given topics. All of them fail to score any marks in the essay as it was difficult for them even to write a sentence correctly.

Situation 2: In a CBSE school in Hyderabad, more than half of the 9th class students fail to score well in the English paper in their final examination. Later, their English teacher finds out that they fail despite having good writing ability because they are slow at writing and thus, cannot complete writing in the stipulated time. Also, they are not given any marks for their excellent speaking skills.

Situation 3: A high school (ICSE) teacher is asked by the school authorities to assess her students in the classroom using her own tasks. She tries to do prepare tasks but fails. So she collects some ready-made questions from guide books available in the market and makes a question paper of it.

Situation 4: A teacher teaching in a private high school (CBSE) demands and gets permission from the school authorities to design his own question papers for mid-term tests. He gives equal weighting to LSRW, Vocabulary and Grammar, allows more time to students who are slow at writing, and includes a small project work in the assessment.

---

2   The researcher discusses the responses of each participant with the whole group and leads them see the advantages of classroom assessment and how teachers' ability to assess plays a major role in it.

Task 6 Group Discussion **(30 minutes)**

Aim: This task helps the participants to have a basic idea about language ability and its assessment, student achievement and the importance of classroom assessment and teachers' ability to carry out assessments in the classroom.

1   The researcher asks the participants to reflect on what they have learnt from the previous five tasks and share with the group what they think about the newly introduced information and knowledge. They are also encouraged by the researcher to ask questions if they have any doubt regarding the same.

2   The researcher asks the participants to reflect on the tasks a little more and maintain a written/electronic/audio account of their reflection and share it with the researcher.

APPENDIX L

Sample of Researcher's Field Notes

Final test — pen — paper — achievement test

- Test of grammar — article, voice, S-V agreement

Blank fill   error   Sentence
              correction   conversion

- time restriction — 20 min's
- Almost discrete point
- No communicative context
- Dictate answers — Absents counted
- Externally calculated but may not help learning
- Spot correction — self-correction
- Very Techni f ⟹ high scores indicators of learning
- Students encouraged explicitly to score well.

Pre

T3 C3 - Tchr extend with answer scripts
↓ (annotated)
- called out eachstudent → returned
- Q-A discussion → answer elicited from each
  ↓   ↓Student
  the probed, were → failed
- Got upset with new students in Dew
  series
- "No practice leads to poor result"
- ∅ 'Tis hard' "my dear ones"

- Explained poor performance   lack of
                                    self motivation
   weak financial bad   ↓
   background   atmosphere   don't
                  at home   pay
                   ↓   attention
                  no encouragement   in class.

- No feedback as such
  students were relieved when she left

- Tchr was at ease, no apprehension abt
  observation. — English
- Topic Any — Class VIII textbook —
- Tchr read aloud / para → a student read aloud
  same para → tchr asked abt comprehension q's to.
  → failed / students answer
  tchr answered → how Telugu to answer
- how Ans repeated → whole para carried
- MT (Telugu) used for explaining meaning of difficult words
- Comprehension q's as homework — gives 2 students
  → all end if learn fairly

  Tchr believed in hard work

- better pronunciation, comprehension, reading ability
- friendly but strict
- Easy and cool.

- Tchr came a lil late
- Told students abt ___ of ___ ___ class

prop 1 Sntml ___

film/book review

- Quietly wrote all the instructions on d bb
  ↓ + took 5 min   format, style, marks
- Tlkd abt each on elaborately
- Asked stdnts if they had undrstnd evrythng
- Clarified doubts
- ___ on spelling/grammatcal errs — plagiarism
  from net
- ___
  —ve marking for evrythng

- Askd stdnts 2 read a classic (book)
  ___ or watch a classic (movie)
  & write a review
- 2 weeks were given
- Stdnts were happy abt it
- No assmnt on their writing

- Very unhappy abt class observation
- Not wrong, a bit irritated

_____ grp. 3 in each

_____ grps are topic - Active/Passive
role of woman in Indian societies

- 10 mins time was given → all grps
- Discussion start after 10 mins
- Strict abt turn-taking
- Each grp puts its arguments

- After each grp presentation → tchr comments on grammatical & communication errors
- Also what they did talk abt during their presentation

- Conclusion - Tchr advises how to improve speaking - some tips

                          ↓
                    pronunciation
                    grammar
                    vocab
                    etc

APPENDIX M

Interview Questions

Pre-Intervention

1 How much time do you spend on assessment?

2 What are the challenges faced by a teacher like you when it comes to assessment?

3 What role does your school play in this?

4 Do you follow CCE?/ Are you aware of changes in assessment policy and CCE?

5 Do you have a manual for internal assessment? How helpful is it?

6 How do you prepare internal assessments?

7 How do you provide feedback to students on their performance?

8 How do you find out about the quality of your assessment tasks?

9 What are your views about training in assessment for all English teachers?

While-Intervention

1 Do you think this training programme will help you in designing internal assessments?

2 Which of the components do you think will help you more?

3 Were the tasks used in today's session difficult for you?

4 Do you have any suggestions to make the tasks used during the training better?

5 Have you started using any of the things you have learnt during the programme?

Post-Intervention

1 What are the things, learnt during the training programme, you can use for internal assessment?

2 How do you plan to use your newly gained knowledge in your school?

3 What changes will you bring in your assessment practice?

4 How will you provide feedback to students?

5 How will you find out whether an assessment task is a good one?

6 Do you have any suggestions to improve the training programme?

7 What kind of training do your colleagues need to carry out internal assessments effectively?