

**Classroom-based assessment and its
relationship with students' self-efficacy:**
The case of English language learning in Rwandan lower
secondary schools

Viateur Ndayishimiye

Doctor of Philosophy
University of York
Education

April, 2018

Abstract

This study was conducted to explore the classroom assessment practices, specifically the forms of assessment and methods of providing feedback used by the teachers of English in lower level Rwandan secondary schools. It also aimed to investigate the students' perceptions of those teachers' assessment practices, the students' self-efficacy for the four English language skills and how such self-efficacy was related to the teachers' assessment practices. Quantitative and qualitative approaches were used to collect and analyse data from both teachers and students. Data was gathered using questionnaires, interviews, focus groups and classroom observations. The obtained data was analysed using SPSS and content analysis.

The results of the study indicated that controlled assessment, commonly known as *paper and pencil* was the most prevalent form that teachers used for assessment. The results also showed that the students reported relatively high positive perceptions for their teachers' classroom assessment practices despite the teachers' use of very limited range of assessment forms and less informative feedback. However, in this high stakes-exam context, it appeared that the students' lack of awareness of the alternative approaches to assessment was responsible for their inability to critically question the teachers' assessment practices; they perceived these as being adequate for the preparation of the summative, high stakes national exam. The results also suggest that students had high self-efficacy in all the English language skills except for listening. The Spearman correlation analysis indicated that the use of some performance assessment correlated with higher levels of students' self-efficacy for productive skills.

These results tend to support findings from previous studies that some assessment practices can affect the students' self-efficacy. They expand the literature and deepen our understanding of the teachers' assessment preferences in English as a Foreign Language (EFL) context and highlight the complexity of the influencing factors of the students' self efficacy. The current findings have major implications for teacher in-service and school-based formal training to raise the teachers' awareness of how important their classroom assessment practices are and how they may affect their students' self-beliefs and learning as a whole.

Table of Contents

Abstract	2
Table of Contents	3
List of Tables	7
List of Figures	9
Acknowledgements	10
Declaration	11
Chapter 1: Introduction	12
1.1 Setting the scene	13
1.2 Education system in Rwanda.....	15
1.3 The Rwandan classroom context	17
1.4 Language education policy in Rwanda.....	18
1.4.1 Multilingual education	19
1.4.2 Prioritisation of English	21
1.4.3 Pressing challenges	23
1.5 Rationale of the study	25
1.6 Purpose and significance of the study.....	27
1.7 Structure of the thesis.....	28
1.8 Summary of the chapter	30
Chapter 2: Literature Review	31
2.0 Introduction.....	31
2.1 Definitions and distinctions	32
2.1.1 Measurement and evaluation	32
2.1.2 Testing.....	33
2.1.3 Assessment.....	39
2.2 The case for effective classroom assessment practices.....	44
2.2.1 A three-element framework	46
2.2.2 The formative purpose of classroom-based assessment.....	48
2.2.3 The role of feedback in classroom-based assessment	49

2.2.4	Forms of classroom-based assessment.....	51
2.2.5	The use of Rubrics in classroom assessment	53
2.3	Recent empirical research findings on classroom assessment practices	55
2.4	Overcoming the challenges in classroom-based assessment	57
2.4.1	Teacher assessment literacy and issues of validity and reliability	58
2.4.2	Influence of contextual factors.....	61
2.5	Classroom assessment and learner self-efficacy	64
2.5.1	Definition and distinction with other motivation constructs.....	65
2.5.2	Sources of self-efficacy.....	67
2.5.3	Effects of self-efficacy on learning motivation and achievement.....	71
2.5.4	Influence of assessment on learner self-efficacy	73
2.6	Summary of the Chapter	77
Chapter 3:	Methodology.....	78
3.0	Introduction.....	78
3.1	A mixed methods approach	78
3.2	Research questions.....	80
3.3	Methods of data collection.....	82
3.3.1	The questionnaire survey	82
3.3.2	Teacher interview.....	90
3.3.3	Student focus groups.....	92
3.3.4	Classroom observation.....	93
3.4	Quality issues: Validity and reliability of research instruments	95
3.5	Research participants	99
3.5.1	Sample size and sampling process	100
3.5.2	Participants in the questionnaire survey.....	106
3.5.3	Participants in the teacher interview	109
3.5.4	Participants in student focus groups.....	111
3.6	Pilot process.....	112
3.7	Data analysis procedures.....	113
3.7.1	Quantitative data analysis	114
3.7.2	Qualitative data analysis	115
3.8	Ethical considerations	118

3.9 Summary of the chapter	121
Chapter 4: The teachers' assessment practices and the students' perceptions	122
4.0 Introduction.....	122
4.1 Forms of assessment used in the classroom.....	123
4.1.1 Paper and pencil versus performance assessment.....	124
4.1.2 The use of rubrics.....	139
4.2 Other assessment practices	141
4.2.1 Purpose of teachers' assessment	142
4.2.2 Time of classroom assessment.....	146
4.2.3 Content focus of assessment	148
4.2.4 Sources of assessment tasks.....	151
4.3 Teachers' methods of providing feedback	153
4.4 Students' perceptions of teachers' assessment practices	160
4.5 Summary of the chapter	163
Chapter 5: Students' self-efficacy and its relationship with teachers' assessment practices	165
5.0 Introduction.....	165
5.1 Students' self-efficacy in the four English language skills.....	166
5.2 The relationship between the students' self-efficacy and the students' perceptions.....	179
5.3 Relationship between the students' self-efficacy and the teachers' assessment practices	182
5.3.1 Relationship between the use of assessment forms and the students' self-efficacy.....	184
5.3.2 The relationship between the use of feedback methods and the students' self-efficacy	188
5.4 Summary of the chapter	190
Chapter 6: Discussion	192
6.0 Introduction.....	192
6.1 Teachers' preferred assessment practices	193
6.1.1 Prevalent use of <i>paper and pencil</i> based forms of assessment	193
6.1.2 Assessment of learning and a focus on linguistic skills	196
6.1.3 Teachers' feedback centred on Marks	201
6.2 Factors influencing the teachers classroom assessment practices.....	205
6.2.1 Scarcity of assessment resources and teachers' limited assessment literacy	206
6.2.2 The Washback effects	209
6.3 The students' perceptions about the teachers' assessment practices	213

6.4 High self-efficacy for reading and speaking	217
6.5 Factors influencing the students' self-efficacy	219
6.5.1 Students' perceptions and their self-efficacy	220
6.5.2 Influence of demographic factors	222
6.5.3 Influence of school context	225
6.5.4 The influence of classroom-based assessment on students' self-efficacy.....	228
6.6 Summary of the chapter	231
Chapter 7: Conclusion.....	233
7.0 Introduction.....	233
7.1 Summary of the current study findings.....	233
7.2 Implications of the study.....	236
7.3 Limitations of the study	238
7.4 Contribution of the study	240
7.5 Suggestions for further research	242
7.6 Personal reflection	244
Appendices.....	247
References.....	289

List of Tables

Table 3.1: Structure of the questionnaire for students	85
Table 3.2: Structure of the questionnaire for teachers	88
Table 3.3: Sample size by province	102
Table 3.4: Gender and class level of the students who participated in the survey	107
Table 3.5: Gender and taught class of teachers who participated in the survey	108
Table 3.6: Teachers' gender by school status.....	108
Table 3.7: Description of the teacher's interview	110
Table 3.8: Description of students' focus groups	111
Table 4.1: Forms of assessment used by teachers in the classroom (Teacher self report)	125
Table 4.2: Students reported forms of assessment used by teachers in the classroom	128
Table 4.3: Teacher-student classroom interaction	132
Table 4.4: Classroom assessment events	136
Table 4.5: Teachers' purpose of assessment.....	143
Table 4.6: Time of Classroom Assessment.....	146
Table 4.7: Content focus of teachers' assessment	149
Table 4.8: Source of teachers' assessment activities	151
Table 4.9: The methods through which students received feedback	153
Table 4.10: The students' perceptions of the teachers' classroom assessment practices	161
Table 5.1: Results from the normality test of the students' self-efficacy for the four skills	167
Table 5.2: Central tendency of the students' self-efficacy for each skill.....	168
Table 5.3: Mann-Whitney Mean Ranks of the students' self-efficacy by gender	169
Table 5.4: Kruskal-Wallis Mean Ranks of the students' self-efficacy by age group	170
Table 5.5: Mann-Whitney Mean Ranks for students' self-efficacy by school learning mode ..	171
Table 5.6: Mann-Whitney Mean Ranks of the students' self-efficacy by location of school ..	172
Table 5.7: Kruskal-Wallis Mean Ranks of the students' self-efficacy by parents' education..	174
Table 5.8: Kruskal-Wallis Mean Ranks of the students' self-efficacy by school status.....	175
Table 5.9: Mann-Whitney Mean Ranks of Students' self-efficacy by teacher's gender	175
Table 5.10: Students self-efficacy (sample from focus group A).....	177

Table 5.11: Spearman’s rho correlation between students’ self-efficacy (SE) and students’ perceptions	179
5.12: Correlation between the teachers’ use of assessment forms related to speaking and the students’ self-efficacy in Speaking (Data from teachers)	184
5.13: Correlation between the teachers’ use of assessment forms related to speaking and the students’ self-efficacy in Speaking (Data from students)	185
5.14: Correlations between the teachers’ use of assessment forms related to writing and the students’ self-efficacy in Writing (Data from teachers)	186
5.15: Correlations between the teachers’ use of assessment forms related to writing and the students’ self-efficacy in Writing (Data from students)	187
5.16: Correlation test between the teachers’ use of feedback methods and the students’ self-efficacy.....	188
5.17: Differences in the students’ self-efficacy based on their perceptions of teacher feedback	189

List of Figures

Figure 1.1: The map of Rwanda (Source: CIA, The World Factbook)	14
Figure 2.1: Recent trends in classroom assessment (from McMillan, 1997).....	45
Figure 2.2 Elements of classroom based assessment practices (Adapted from Hill, 2017).....	48
Figure 2.3: A model of feedback to enhance learning (from Hattie and Timperly, 2007)	50
Figure 2.4: Bandura's model of reciprocal interactions (from Schunk and Zimmerman, 2007)	65
Figure 2.5: the role of self-efficacy in academic achievement (adapted from Schunk, 1995) ...	72
Figure 2.6: Effect of classroom assessment on the student's achievement by mediation of student's self-efficacy (adapted from Ross et al., 2002).....	74
Figure 3.1: Research questions and methods of data collection	81
Figure 3.2: School Learning mode (n=41) Figure 3.3: School Status (n=41)	103
Figure 3.4: Location of the surveyed schools (n=41)	104
Figure 3.5: Students' age by gender (n=1242).....	104
Figure 3.6: Students' parents education by school learning mode	105
Figure 3.7: Teacher in-service training in a 5 year period (n=66)	106
Figure 3.8: Example of first cycle analysis of teacher interview using nVivo coding	116
Figure 3.9: Sample of codes categorisation (teacher interview).....	117
Figure 4.1: Frequency of use of Paper and pencil vs performance based assessment forms ...	126

Acknowledgements

My thanks and gratitude go to my supervisors Dr. Bill Soden and Prof. Chris Kyriacou for their guidance and encouragement throughout the study. Special thanks go to Dr Bill for his invaluable feedback that helped me to find a way forward whenever I felt like there was none. His clear and timely comments gave me motivation to continue and subsequently complete the programme. I would also like to thank the staff from the Department of Education, the Math Skills Centre and the library for the assistance that they provided to me in one way or another.

I am deeply grateful to the Commonwealth Scholarship Commission for granting me the scholarship to undertake my studies at the University of York. Without their financial support, this study would not have been possible. I am also thankful to Rwanda Education Board and to the management of the College of Medicine and Health Sciences at the University of Rwanda for their endorsement of my candidacy to a PhD research programme in the UK. Their support was of great significance and it sustained my commitment and dedication to the programme.

I also wish to thank all those who participated in the study particularly all the secondary school teachers and their students who took time off their busy schedules to take part in the data collection process. Their insights and perspectives were indispensable. I am also thankful to Dr Telima Adolphus for his unreserved participation in the proofreading process. His feedback and comments were highly invaluable.

Finally, my heartfelt appreciation goes to my wife Jeannine and my daughter Briella for their unconditional love and moral support to me. Their company and total care made me feel at home in the UK while away from friends and the rest of my family. I humbly dedicate this work to you.

Declaration

I hereby declare that I am the sole author of this thesis. It is a presentation of my original work and it has not previously been presented for an award at this, or any other University. All sources are acknowledged as References.

Chapter 1: Introduction

Classroom assessment has become a popular topic in the EFL context and in the educational research realm as a whole. Broadly defined as classroom activities that teachers use to collect information about learning, adjust their teaching and to improve the students learning (Black and Wiliam, 1998b; Hill and McNamara, 2012), classroom assessment has increasingly established itself as a major component of the teaching and learning process. A number of research studies on EFL classroom assessment have been conducted resulting in significant findings that have expanded our understanding of the classroom assessment practice in the language classroom (Cheng, Rogers and Wang, 2008; Davison, 2004; Mekonnen, 2014; Yang, 2008). In particular, a large amount of literature has advocated the use of assessment activities that give every learner the opportunity to practice and measure their knowledge and skills. Emphasis has been put on formative assessment that aims to empower learners by providing them with useful information that helps them to know where they are in the learning process and what they need to do to successfully reach their learning targets (Black and Wiliam, 2009; Hattie and Timperley, 2007; Rea-Dickins, 2001, 2007).

In addition, current research literature demonstrates that classroom assessment plays a significant role in shaping the students' affective domain and hence affecting the students' motivation (e.g Cheng, Wu and Bettney, 2014; Kang'ethe, 2014; Stefanou and Parkes, 2003; Zimbicki, 2007). During their participation in assessment tasks, students develop perceptions about assessment activities which in turn impact on their motivation. "As students actively process assessment events they develop cognitions concerning task importance or value, difficulty, and the likelihood of success. These beliefs in turn influence expectations, effort and motivation." (McMillan and Workman, 1998, p. 17). In this regard, studies have investigated the link between classroom assessment and the students' motivation commonly viewed as a major promoter of learning (e.g. McMillan and Workman, 1998; Stefanou and Parkes, 2003; Zimbicki, 2007). In the current study, the focus remains on the classroom assessment practices and the students' perceptions of those practices. However, the study goes further to explore the relationship between the classroom assessment practices and the students' self-efficacy in the

context of English as a foreign language classroom. The role of learner's self-efficacy i.e. the learner's beliefs that she or he can successfully perform a specific task, in enhancing the learner's motivation has been widely acknowledged. Although researchers have increasingly been interested in investigating the effects of self-efficacy on academic achievement (e.g Mills, Pajares and Herron, 2007; Pajares, 2003; Schunck, 1982), very few studies have explored factors influencing the learners' self-efficacy. The current study was conducted to contribute in this area by examining the relationship between classroom assessment practices and students' self-efficacy in an EFL context. More discussion on self-efficacy and its role in promoting learning is provided in Section 2.5 of chapter two.

In this chapter, an introduction to the context of the current study is provided by highlighting the key issues in the current language teaching and learning practices in Rwanda. The Rwandan language education context discussed in this chapter dates back to the 1960s, the year that the country gained independence until the date of data collection for the current research project in 2016. A particular focus is put on the period after 1994 when the country experienced drastic socio-political changes and when English became one of the official languages in the country and eventually became the most dominantly used in public administration and in education. The chapter closes by briefly relating the major research findings in language assessment in general and classroom based assessment in EFL contexts in particular to language education and assessment in the context of Rwanda. This in turn allows for a presentation of the rationale of the current study vis-a-vis the presented background. The structure of the study and the main contents in each chapter are also presented towards the end of this chapter.

1.1 Setting the scene

The current study was conducted in Rwanda, a country located in central Eastern Africa. Its total size is estimated at 26,338 square kilometres bordering four countries namely Burundi in the south, Democratic Republic of the Congo in the west, Uganda in the north and Tanzania in the east (see Figure 1.1). As a landlocked country, its economy is based on tourism and agricultural exports mainly tea and coffee. According to the 2012 general census, Rwanda had

a population of nearly eleven million expected to double by 2047 (Ntawitonda, 2014). The country has made significant progress in raising the literacy rate in the last two decades (with 20% increase in the population literacy rate), from 50% in 1995 to 70% in 2015 (CIA World Factbook, n.d).



Figure 1.1: The map of Rwanda (Source: CIA, The World Factbook)

For the last two decades, Rwanda has been widely known for its civil war and genocide against the Tutsis in 1994. This tragic past was a result of a buildup of social and political tensions since the end of absolute monarchy in 1959 and the establishment of the republic system in 1962. The three years between 1959 and 1962 saw many Rwandans from the Tutsi ethnic group flee persecution to the neighboring countries. These would later start an armed struggle to return to their home country in 1990s which ended in 1994, when the former refugees were repatriated. Under the new leadership, the country has made significant recovery both economically and socially. It has been commended for its satisfactory achievement of most of the Millennium Development Goals (MDGs) in 2015 through implementation of programmes that contributed to the improvement of living standards, reduction of poverty and increase in near-universal primary school enrollment (World Bank, 2017)

1.2 Education system in Rwanda

Education in Rwanda is managed by the Ministry of Education (MINEDUC) and its affiliated institutions notably the Rwanda Education Board (REB) and the Higher Education Council (HEC). The Rwandan education system is based on three levels of schooling: primary, secondary and tertiary level. Children aged seven years are expected to start the primary school which lasts six years. Following the language policy changes in 2009, the pupils at primary school are taught in their native language Kinyarwanda from primary one (P1) to primary three (P3) after which English is ideally used as the medium of instruction until P6 (MINEDUC, 2015a). There are also private primary schools that teach in either French or English and either of these two languages taken as an elective subject (Ep Nuffic, 2015). The primary school ends with a high stakes national examination that yields admission to lower/ordinary school. The failing students who choose to continue their education have either to retake the test or join the increasingly expensive private schools. As of 2016, the enrollment rate to lower secondary school was 22.6% of all children with the age of starting the secondary school (MINEDUC, 2016).

The students at the lower secondary school are generally aged between thirteen and fifteen years (MINEDUC, 2016). Since 2010, a new system was launched that combined the three years of lower secondary schools with the six years of primary school to form the fee-free Nine Year Basic Education, commonly Known as 9YBE (MINEDUC, 2010; Pearson, 2013). As a result, some primary schools were provided with the new premises and other facilities to accommodate the large number of enrolling students. The 9YBE schools are usually attended by local students from the school surrounding community who did not perform highly enough in the primary leaving examination to be admitted in the boarding schools. The 9YBE was established to help with the increasingly high cost of attending the highly competitive boarding school that often receives highly performing students. In Rwanda, the secondary schools can either be boarding or day based. The latter include the 9YBE and the twelve year basic education day schools simply known as “Twelve” (12YBE) (Pearson, 2014). These are secondary schools that usually have upper level of secondary but which, like 9YBE, are day schools and usually receives students from communities around the school.

On the other hand, while most of day schools (9YBE and 12YBE) are fully public, the boarding secondary schools are usually long established private schools that are also government aided. Boarding schools also usually possess more facilities than day schools and usually pay higher wages compared to day schools (Williams, 2015). However, all these schools have the same curriculum delivered by MINEDUC. As in the upper primary school, the medium of instruction at all secondary schools is English except for a few private schools that teach in French where English is also taught as an elective subject. At the end of secondary education, students have to take a high stakes national examination test to be awarded the advanced diploma in secondary education. This also yields eligibility for admission to university education for highly performing candidates while others may choose to enroll in Technical and Vocation Education and Training (TVET). These are schools that “provide young people and the unemployed with the skills to gain productive employment and those already in employment with an opportunity to upgrade their skills, including entrepreneurs and those wishing to work for themselves” (MINEDUC, 2015a, p. 9). They include Integrated Polytechnic Schools (IPRC) where students can take courses of up to four years for the award of a bachelor’s degree. Parallel to upper secondary school are also the Vocational Training Centres (VTC) which are part of TVET and that were established to offer both short courses (lasting no more than twelve months) or longer courses that can take up to three years after which a certificate or diploma is awarded (MINEDUC, 2015a). Both TVET and VTC programmes have to be taught in English like other educational sectors in Rwanda.

The highly performing secondary school graduates are admitted to university once they get the cut off score for admission at higher education. However, the number of students admitted to the tertiary level is still low but steadily increasing as the number of universities increases. As of 2011, only 12.7% of those who sat the secondary leaving exam test were admitted to the public universities (Kwizera, 2011). The university education in Rwanda operates on a four years basis i.e. bachelor's' degree with some programmes offering three year courses for the award of diplomas. The language of instruction at university is English with the exception of a few higher learning institutions offering courses in French. Recent changes in public higher education have led to the combination of all public higher learning institutions to create one national university known as University of Rwanda. This was done to “transform the

country' s higher education system by increasing assets, promoting equality, ensuring high quality of education and providing infrastructure of advanced quality” (Ndahiro, Gatere and Nkusi, 2015, p.210). The reorganisation of the higher learning institutions under one management was expected to advance the quality of research in the university and improve the quality of teaching and learning together with designing programs that provide the citizens with the necessary knowledge and skills for national and global competitiveness.

The current study was focused on the lower level of secondary school. The investigation targeted the three initial years of secondary education after P6 where English is taught as one of the core subjects on the curriculum. Specifically, the students in second form (F2) and third form (F3) of lower secondary schools and their teachers were the only participants of the study (See Section 3.5). The participants were students and teachers from 9YBE and 12YBE schools which were both day schools. Some participants were also from boarding schools.

1.3 The Rwandan classroom context

This section looks at the context in which the English language is taught in the Rwandan secondary schools. It describes the English language classroom in the context of Rwanda highlighting the teachers' working conditions with a particular focus on the contextual aspects that may have direct or indirect impact on the teachers' day-to-day teaching practice including their way of assessing students in the classroom. This is important as the contextual factors can also have impact on the learning process as a whole.

The current study targeted schools in which both female and male students were enrolled at an approximately equal rate. This was intended to reflect the general reality of secondary schools in Rwanda which usually enroll both girls and boys. All forty one surveyed schools were mixed schools and they were taught by male or female teachers. In most cases, the number of male students in the classroom was higher than female students. At each school, English was taught by one, two or more teachers depending on the number of the size of the school and the number of used classrooms. Working on a rotation basis i.e. different subject teachers taking turns to teach different subjects to the same student groups, teachers of English

were often assigned to teach designated groups of students for a maximum of up to forty-five 'periods' per week on a full time contract. One 'period' equals to forty minute long session of face to face teaching. The teaching facilities in the classroom were the blackboard and chalk boxes. No other facilities such as audio visual equipment were visibly available. The teachers planned and administered the teaching and assessment activities in the classroom mainly using paper and pen. It is important to note that no specific time officially allocated for other activities such as preparation and marking. Teachers are expected to find extra time beside their regular periods to look for resources and mark students' assignments as they strive to cover the curriculum in preparation of the high stakes national examination at the end of the third form of secondary school. The scarcity of teaching resources and of assessment materials in particular compels many teachers to engage in the constant search of resources which adds more work to their already tight workload.

In many schools, the physical aspect of the classroom did not seem to facilitate teachers in their day to day teaching activities. This is particularly with regard to the class size, the type of desks used and the seats arrangement which all seem to restrict the free movement of both the teacher and the students in the classroom. In 2015, the yearly educational statistics published by the ministry of education showed that the student-teacher ratio in secondary school was 35:1 (MINEDUC, 2015b). It was observed during data collection that some schools had classroom with up to forty nine students. Beside the large class size in some schools, the large, heavy desks used in many school may present a challenge especially for easy communication and interaction between the teacher and the students and among students themselves. Although the rooms are often big with many and usually wide windows, there is often no enough desks. It was also noted that students' books were insufficient or not available at some schools which become another challenge when it came to assigning classroom activities.

1.4 Language education policy in Rwanda

Rwanda is predominantly a monolingual society. Its sole national language Kinyarwanda is spoken by more than 99% of Rwandans and it is used as the language of instruction at nurseries and primary levels of education (MINECOFIN, 2005). It is also used nationwide as a language

of ordinary citizens' commerce and social events and it is spoken at work places throughout the country. Kinyarwanda is also one of the Rwandan official languages in addition to English and French. Granted in the 2003 constitution amended in 2015, the official status of these three languages meant that their use was legal and that "official documents may be either in one, two or all of the official languages" (Article 8 of Rwanda Constitution, Official Gazette Special Number, December, 2015). The post 1994 war and genocide era was thus marked by the political will to implement a multilingual language policy where, beside Kinyarwanda, French and English were both almost equally used in all educational levels after P3.

1.4.1 Multilingual education

Rwanda was one of many African countries where French was used as an official language since the beginning of the colonial rule in early 20th century. French and Kinyarwanda were the only two official languages in Rwanda until 1994 when English became the second foreign language officially recognised in the country. French continued to be used in Rwanda after independence in 1962 and was consolidated by the then growing ties with France in 1970s. Rwanda also joined the French speaking clubs including mainly the Organisation Internationale de la Francophonie in 1970 and was one of the founding members of the Economic Community of Central African States (ECGL) in 1983 (JeuneAfrique, 2016). However, the repatriation of English speaking refugees into Rwanda in 1994 marked the beginning of a new linguistic landscape that the country was about to embrace. English was introduced in schools first as a taught subject from P4 and became one of the most important subjects in the ordinary level curriculum. French and English also had the highest number of hours of instruction per week. In addition, the ministry of education accredited new schools that used English only as a medium of instruction but also encouraged a multilingual based education where the use of English and French were allowed at secondary school level. Article 34 of the 2003 law establishing the organisation and the functioning of nursery, primary and secondary schools stipulated that:

The language of instruction in the first cycle of primary education is Kinyarwanda except for the lessons of foreign languages. The Minister having education in his or her portfolio may, through a Ministerial Order, authorise the use of French or English as the medium of instruction in the first cycle. The language of instruction in the second cycle is French or English, except for other language lessons (MINEDUC, 2003, Article 34).

Overall, French continued to be used as a medium of instruction in the majority of secondary schools as a result of its long establishment in education. As of 2009, up to 95% of all schools from the upper primary level taught in French (McGreal, 2009). However, with the adoption of English as a medium of instruction in many schools and the increasingly frequent use of English in public administration (See Section 1.4.2 on prioritisation of English), the teaching of French and its use as a medium of instruction gradually declined.

Many Rwandans appreciated the evident political will to build a human resource based economy with its multilingual citizens being able to take advantages of the current globalisation. Raising the Rwandans' proficiency in regional languages such as Swahili and in French and English as international languages was expected to help Rwandans seek and get the best of many opportunities available both at home and away as Simpson and Muvunyi (2012) highlight:

Given that the government places a high premium on the development of human capital with the necessary knowledge and skills as a vehicle for socio-economic development – and in line with the Constitution, which stipulates that Kinyarwanda, French and English are official languages – it has developed a trilingual education policy so as to gain regional and international advantages associated with trade, foreign relations, employment and education. (p. 154)

In addition, Swahili was at the same time acknowledged as an important regional language. It was taught in some schools as an optional subject and as a core subject for some language majors. This changed in 2017 when the Rwandan parliament passed the law to recognise Swahili as a fourth official language and was expected to be introduced in the secondary school curriculum as a compulsory subject (Mugisha, 2017).

While some Rwandans expressed their support to the introduction of these regional and international languages in schools, the recurrent and unproductive aspect of these changes has been viewed by some as an indication of the lack of a stable language policy. This started to raise concerns among many in the academia, research elite and teachers alike particularly since the introduction of English as the language of instruction while the majority of teachers spoke French. This move was viewed by many as “sudden and mandatory, and the reasons for it political and economic” (Pearson, 2014, p. 43). It appeared that the priority was almost

exclusively for English to be the dominantly used language beside Kinyarwanda native language.

1.4. 2 Prioritisation of English

As highlighted above, English increasingly got higher leverage over French since 1994. The 2003 constitutional law that made English the official language and subsequent changes in language education policy both consolidated the importance that English was increasingly occupying in the education sector. In fact, the mass repatriation of English speaking former Rwandan refugees from English speaking countries of Uganda and Tanzania was an important factor behind the subsequent prevalent use of English language over French in administrative and educational settings (McGreal, 2009). Unlike the French speaking Rwandans who started to learn English, the English speaking Rwandan repatriates demonstrated little interest in learning French. Instead, they became a pivotal force behind the promotion of English alongside French and Kinyarwanda as Samuelson and Freeman (2010) write:

Because members of the Rwandan elite are often former refugees who grew up in Uganda, English stands for political power. This was particularly so after 1994, when the newly arrived Anglophones established English as an official language and made little attempt to learn French. These returnees are now the establishment. (p. 202)

The change of the language of instruction and other English promoting policies were claimed to be driven by the need for the regional and international integration (Basheija, 2014; Tabaro, 2012). Rwanda joined the East African Community (EAC) in 2007 and became a member of Commonwealth of Nations in 2009 both of which use English as their official and working language. It is also important to note that these new geopolitical alliances came at the time of weakening relations with France in 2006 which Rwanda accused of having provided support to the perpetrators of genocide against the Tutsis in 1994 (McGreal, 2009).

Nonetheless, the prioritisation of English and evidence of its important status became more and more visible particularly in public institutions despite the discontent of some Rwandans that French was unlawfully deprived of its constitutional status. In 2014, one political party made an unsuccessful attempt to block the removal of French on the newly

released bank notes (Dusabemungu, 2014). In addition, although it is permitted and granted by the constitution law to use any of the four official languages, the use of English even in situations where all communicators spoke the native language was not uncommon. These include interactions during ordinary meetings or emails exchange between co-workers. As a result, many public servants who were fluent speakers of French started to join the newly established private centres to learn English particularly for their everyday communication in the workplace.

In education, special effort was put on training teachers in order to increase their proficiency in English and to be able to use it in the classroom. Since 2009 when the cabinet voted for a law instituting English as a sole medium of instruction from P4 level, several programmes have been put in place by the ministry of education through REB to raise the English language proficiency for primary and secondary school teachers so that they can be able to teach their subject in English effectively. This was a huge undertaking given that very few teachers had studied English before. In actual fact, when the law was passed to shift the medium of instruction in 2009, only 4,700 of 31,000 primary schools teachers (15%) and 600 of 12,000 secondary school teachers (5%) had had English courses in their previous education (McGreal, 2009). Despite the apparent challenges that the implementation of the new policy was likely going to face, all the public schools started implementing it in support of the ministry and its partners mainly the British Council Rwanda.

In 2012, the Ministry of Education launched the School-based Mentorship programme (SBM) in 2012. Through SBM, the ministry hired local and regional language experts from the neighbouring countries mainly from Uganda to train primary and secondary school teachers of English and other subjects on the methodologies to use in the classroom. These school based mentors were also supported by senior mentors who were trained by the ministry in collaboration with the British Council Rwanda (REB, 2014). The school-based mentorship programme was later called off after four years of its inception under reports of mismanagement and inefficiency of the programme (Kanamugire, 2017). However, the Rwanda Education Board praised the contribution that SBM rendered to the development of the teachers' skills in using English as medium of instruction. A new scheme was initiated to

replace SBM where local teachers who were fluent in English were designated and provided training to help their colleagues who still struggled with teaching in English. In addition to helping with the training of mentors, the British Council Rwanda, in partnership with the local Association of Teachers of English in Rwanda (ATER) and the International Educational Exchange (IEE), implemented the “Supporting Teachers' English through Mentoring (STEM)” project through which teachers from 36 primary schools were provided with self-study materials and training to help them improve their English proficiency and teaching skills (British Council Rwanda, 2015).

English is currently assessed at the classroom level with teacher-made tests often for promotional and administrative purposes. It is also assessed at the national level usually for promotional and certification purposes for students leaving the lower secondary schools and at the end of year six of secondary education respectively. Tests at the national level are the most important and high stakes assessment prepared by Rwanda national examination council (RNEC). In early 2018, the ministry of education announced its plan to introduce English language as a core subject in all public universities (Mucunguzi, 2018). Students will be expected to demonstrate a satisfactory level of English throughout their faculty studies and will be required to take a compulsory English test before they could be awarded degrees.

1.4.3 Pressing challenges

The teaching and learning of English language in Rwanda has been an educational priority for the last two decades and significant progress has been unquestionably made in promoting the use of English in the country. English has become the commonly used language in public institutional events beside Kinyarwanda and is predominantly used in official documents. Its use as a medium of instruction has also motivated Rwandans in the education sector including teachers in secondary schools to improve their proficiency and become more confident in using it in the classroom as a medium of instruction (Odeke, 2015). However, there is still more to do to raise the learners' proficiency. It has often been reported that the students' level of English in Rwanda is still low despite the aforementioned efforts that have been put in place since the 1990s. Graduates at all levels of education in the country are said to demonstrate high difficulties in speaking English (Rwirahira, 2017; Odeke, 2015; Osae, 2015). These also

include university graduates who fail to communicate in English effectively once at workplace (Tabaro, 2012).

The contextual realities highlighted in the above section particularly the shortage of appropriate resources and the teachers' low level of English that limits their interaction with students during instruction are some of the commonly cited challenges that slow down the implementation on the language programmes (Reddick, 2015; UNESCO, 2014). Due to their low proficiency in English, teachers prefer to limit their interaction with students and restrict their teaching to the "chalk-and-talk structure" (UNESCO, 2014). This in turn limits the opportunities to students to practice the language in the classroom, which is the only place for many Rwandan students to speak English. The fact that Rwandans have one common national language spoken across the entire country often leads to the common "fear of being seen as arrogant" once they speak another language other than Kinyarwanda (Tabaro, 2012, Par. 3). Using English in the classroom therefore seems to be important for many students who have little or no exposure at all to English when they are not in school. The relevance of speaking English in schools to the learners' proficiency development was recognized by REB in 2012, when it urged all secondary school teachers and students to use English only for communication at school. This was another effort to increase language exposure and to encourage the students and teachers to develop their communicative skills through practice (Uwishyaka, 2015).

Using the target language in the classroom is likely to help students "perceive it as a useful medium for communication and develop more positive motivation to learn it" (Littlewood and Yu, 2011, p. 66). It appears that in a monolingual context like Rwanda, engaging learners in real-life like activities in the language classroom can be fundamental for the learners' target language development. As highlighted in the following sections, the focus of the current study was on assessment activities that were used in English language classroom and the extent to which it influenced the students' perceptions and self-efficacy for communicating in English.

1.5 Rationale of the study

Given the context described above, the current study has been conducted to investigate the current issues in English language teaching and learning in Rwandan lower level of secondary schools. My personal experience as a former teacher of English in Rwandan secondary schools raised my interest in investigating the classroom assessment and the students' self-efficacy to communicate in English. This followed my observation that there was little research done regarding the classroom assessment practices in the Rwandan context and the apparent paucity of relevant studies that dealt with the relationship between classroom assessment and the students' self-efficacy in EFL context in general.

During my four years of English teaching experience at the University of Rwanda, I realised that the students exhibited a high level of anxiety for using English inside as well as outside the classroom. This increased my interest to investigate what was behind their anxiety and their demonstrated limited enthusiasm to communicate in English. I learned from my casual talks with some of these students that they felt less capable of speaking English accurately and fluently and therefore preferred to use their mother tongue or limit their participation in the classroom whenever possible. Most students attributed their lack of confidence to their previous English learning experience at secondary schools reporting that it was ineffective because the teaching and assessment was solely focused on grammar. This caught my attention and I started searching and reviewing literature on English language teaching and assessment practices in Rwandan secondary schools to understand the assessment practices of English teachers and ultimately examine whether and the extent to which such practices were related to the students' self-efficacy for communicating in English.

However, I quickly noticed the paucity of literature on classroom assessment practices in secondary school in Rwanda. There appeared to be no single study available on the topic of students' self-efficacy in Rwanda and research on English language teaching and learning was particularly new partly due to the linguistic background of the country where French and Kinyarwanda had been the only two languages widely used in Education (See section 1.4). In addition to the lack of literature related to the Rwandan context, I also realised that there was

very limited research on the assessment practices of teachers of English in EFL context in general. Also, while some theoretical works and research studies have explored the link between the classroom assessment and students' motivation (e.g Alkharusi, 2013; Baadte and Schnotz, 2013; Harlem and Crick; 2002), there is limited literature on the relationship between classroom assessment practices and the learners' self-efficacy as a motivational construct (e.g Zimbicki, 2007). Although some literature has begun to emerge on the effects of classroom assessment on learner self-efficacy in science (e.g Alkharusi et al. , 2014, Dorman, Fisher and Waldrip, 2006), there is little research in EFL in general that has explored the relationship between what teachers do during the classroom assessment and the students' self-efficacy.

Overall, there is an apparent need to investigate further whether and to what extent what happens in the EFL classroom in terms of assessment is related to the learners' self-efficacy. Kuciel (2013) comments on this relationship between assessment and self-efficacy stressing that "As self-efficacy is postulated to be chiefly the product of mastery experience, the students' perceptions or experiences of previous performance in the foreign language can be crucial for establishing high self-efficacy." (p. 37). In this regard, it appears that the use of activities that mirror the real life use of the language during assessment offers an important experience to learners and can therefore be a way of enhancing their confidence that they can perform tasks successfully (self-efficacy). As McMillan and Workman (1998) highlight, assessment practices can help foster the students' self-efficacy that in turn increases the learning motivation. The current study was conducted to add a research contribution to the literature in this research area by examining the classroom assessment practices of EL teachers and the EFL learners' self-efficacy while relating it to the Rwandan context.

In addition to the scarcity of research in EFL classroom assessment and its relationship to learners' self-efficacy, it appears that the studies done in this area often used the quantitative methods, collecting data mainly by use of questionnaire surveys. Most of the current literature on classroom assessment and self-efficacy is derived from studies that have often used quantitative methods for both the collection and analysis of data (e.g. Beleghezadeh and Masoun, 2013; Deluka et al. 2016; Dunn, Strafford and Marston, 2003; Inbar-Lourie and Donitsa, 2009; Jenks, 2004; Tzagari and Pavlou, 2009; Yang, 2008). Some studies have also

relied on qualitative methods (e.g Cheng et al. 2004; Chumun, 2002) with a few others have using mixed methods (e.g. Mekonnen, 2014; Mussawy, 2009; Zimbicki, 2007). While the findings from these studies have made significant contribution to the current literature, it is also important to have research findings derived from data collected and analysed by use of mixed approach. This helps to avoid weaknesses inherent in either methods and allows for data triangulation. It is in this regard that the current study used both quantitative and qualitative methods to add a methodology contribution to the current research in the topic of classroom assessment and self-efficacy.

1.6 Purpose and significance of the study

This study sought to investigate the classroom assessment in the Rwandan lower secondary school as it is in this education that students have the most time to learn English and where English has the most hours on the teaching timetable. The students are thus expected to get the solid foundation of language competence in preparation for their studies in upper secondary school and beyond. The investigation was conducted under the title of “classroom based assessment and its relationship with the students self-efficacy”. This was intended to explore the teachers’ assessment practices and the student's self-efficacy as a motivational factor for communicating in English. In other words, the study sought to explore the students’ self-efficacy in the four English language skills and to determine the extent to which their self-efficacy was related to the teachers’ classroom assessment practices. Specifically, the study aimed to investigate:

1. The forms of assessment and methods of providing feedback that teachers of English used in their English language classroom at the lower level of secondary schools in Rwanda
2. The teachers’ purpose and time of assessment as well as content focus and source of their assessment activities
3. The students’ perceptions of their teachers’ classroom assessment practices
4. The students’ self-efficacy for the four English language skills, i.e. speaking, writing, listening and reading
5. The relationship between the students’ self-efficacy in the four English language skills and the teachers’ classroom assessment practices

The results of the study were expected to shed light on the practice of classroom assessment in Rwanda as an EFL context and to explore the role of assessment to the development of EFL learners' self-beliefs. The key beneficiaries of the study findings are the educational research community in Rwanda education in particular and in EFL discipline in general. The research in the areas suggested herein could lead to further and deeper understanding of the EFL teachers' assessment practice and the students' self-efficacy for communicating in English. The teachers and the students in lower level of secondary schools are also the direct beneficiaries of the current study results. The findings can be used to raise the teachers and students awareness on how assessment can affect learning as a whole. Classroom assessment can be the source of information for language learners about their abilities to perform communicative tasks. During the assessment process, students get the opportunity to practice and to be evaluated on different tasks where the feedback that they get from peers and teachers also develop their awareness of their strengths and weaknesses. The results of this study can also help to raise the teachers' and students' assessment literacy and to help calibrate the students' judgment of their self-efficacy. The realistic beliefs about their capabilities to perform tasks in English are expected to lead to the adoption of the right strategies for improved learning and use of English language. The results of this study could also be of interest to education policy makers in Rwanda. Its findings could serve as a source of information on the implementation of language programmes at the lower level of Rwandan secondary school system.

1.7 Structure of the thesis

The thesis is comprised of seven chapters with each chapter focusing on one specific issue. Following the current introduction, chapter two discusses the theoretical background of the study. It reviews the conceptual and research underpinnings in areas of classroom assessment in education as a whole and in the EFL context in particular. Chapter two also provides an overview of the current literature on academic self-efficacy with emphasis on EFL context and highlights the most recent findings about the role of self-efficacy in language learning. Chapter three follows with a detailed account of the methodological approach that informed the design of data collection instruments and guided the process of data collection. Information on

relevant key methodological issues such as sample and sampling techniques and data analysis methods is also provided in chapter three.

Chapter four and chapter five are centred on the presentation of the results from the analysis of data on the teachers' classroom assessment practices and data on the students' self-efficacy respectively. In chapter four, the focus is on the results from the analysis of data gathered to answer research question one and its sub questions regarding the assessment forms that teachers used in their everyday classroom assessment, the methods that teachers used to provide feedback to students, other assessment practices and the students' own perceptions of the teachers assessment practices as a whole. Chapter five deals with the results from the analysis of data on research question two about the students' self-efficacy levels in the four English language skills (speaking, writing, reading and listening). The results presented in these two chapters relate to quantitative and qualitative data analysed side by side for ease of interpretation. These results are discussed further in chapter six. The discussion in chapter six revisits the study major findings and identifies observable key patterns and relationships shown by those findings. The chapter also refers to the previous literature not only to put the findings in a wider perspective but also to highlight further the significance of the results in the context of the current study.

The thesis closes on some notes about the study contribution, its limitations, implications and suggestions for future research. These are briefly discussed in chapter seven. The chapter opens with a summary of the findings and goes on to describe the educational and research implications from the major findings. It also explains how the current research supports the broader knowledge and understanding of the practice of classroom assessment and the construct of self-efficacy in EFL settings. Some limitations of the study are also listed and their importance in relation to the interpretation of the results is discussed. The chapter ends with a suggestion of further research followed by my personal reflection on the study.

1.8 Summary of the chapter

This introductory chapter has focused on the study setting and the circumstances that led to the current study. It is clarified that the study was set to investigate the teachers' classroom assessment practices of the lower level of secondary schools in Rwanda following the introduction of English as a major curriculum subject and as a language of instruction since 2009. It is argued that in spite of the far reaching implications of this new language policy, there seemed to be insufficient research to investigate the implementation of English language teaching and assessment especially at the lower level of secondary school that has the most hours allotted for English as a curriculum subject. There is a warrant for more studies especially as some evidence point to the limited English language proficiency on the side of Rwandan learners and a demonstrated lack of confidence in terms of English language use both in the classroom and outside the school. In this regard, it is argued in the chapter that there is still a need for more literature about the EFL learners' confidence and self-belief for using English. The chapter also highlights the significance of the current findings and identifies the different ways in which the results can contribute in bringing about improvement in the assessment of English in Rwanda for better learning and use of English among Rwandan learners. The study also adds to the growing body of literature on self-efficacy and classroom assessment in EFL and the use of different methods of data collection and analysis offers another evidence of the importance of using mixed methods. The findings can also serve as a source of information for teachers to know how their classroom assessment practices can be influential in the development of the students' beliefs and how they can help the students calibrate their capability judgement.

Chapter 2: Literature Review

2.0 Introduction

This chapter reviews the concepts and the theoretical perspectives relating to the topic of the current study. As this study sought to investigate the teachers' classroom assessment and its relationship with the students' self-efficacy for communicating in English, the review focuses primarily on literature related to the classroom assessment and self-efficacy in general and in EFL context in particular. The chapter opens with a presentation of definitions and distinctions of the concept of assessment in education in general and in EFL assessment in particular. Given the important role that testing assumes in assessment in general and its relevance to the topic of the current study in particular, the section also discusses some of the major trends that have characterized the language testing field. This is discussed in order to understand how language testing has evolved overtime and to highlight how it has influenced the classroom-based assessment. This discussion illustrates various changes and shifts in language testing paradigms over the last few decades, from the psychometric-structural linguistics period to the current performance oriented trends.

Central to this chapter is the presentation of the case for the classroom-based assessment which is described as an alternative to the overwhelming reliance on standardized testing. A concise review of the findings from previous empirical studies is also presented to draw attention to the current theoretical and practical issues that have caught interest of researchers so far as well as suggested solutions to challenges that arise during the implementation of classroom assessment. The review focuses on relevant educational research and EFL studies that have contributed to the understanding of the classroom assessment and its impact on learning. In addition, the current chapter discusses theoretical works and research findings on learner self-efficacy particularly in EFL contexts. Defined as "People's' judgment of their capabilities to organise and execute courses of action required to attain designated types of performances" (Bandura, 1986, p. 391), Self-efficacy in academic settings has been viewed as a source of motivation to students to take up more challenging activities leading to high

achievement (Pajares, 1996; 2003). The review of literature on self-efficacy also puts special consideration to studies that relate the classroom assessment to the learner self-efficacy. Although self-efficacy as a motivational aspect is relatively new in EFL research, an increasing number of research studies have provided important literature on the role of EFL learners' self-efficacy in their language learning achievement.

2.1 Definitions and distinctions

The conceptualization of classroom assessment entails first of all the elucidation of differences existing between assessment and other related terminologies. These terms include *measurement*, *evaluation* and *testing* which are commonly used in the educational assessment realm and sometimes used interchangeably with *assessment*. The following sections briefly define these terms and a special attention is given to *assessment* and *testing* as they are central to the current study.

2.1.1 Measurement and evaluation

Assessment is sometimes used interchangeably with *measurement*, a term used in educational settings to refer to the process of determining the degree to which a student performs towards satisfaction of preset standards or objectives. Measurement is defined as “the process of quantifying the observed performance of classroom learners” (Brown and Abeywickrama 2010, p.4). Teachers are involved in the measurement process in their classroom assessment practice when they score quizzes and tests by assigning a numerical score to the students works (Airasian, 2005). It is worth mentioning however, that the ‘quantification’ process referred to here is not necessarily about assigning quantifiable measurements to students’ performance. Bachman (1990) explains that qualitative forms such as written descriptions and oral feedback are non quantifiable measurements which can also be used for measurement purposes. Measurement can be essential to the learning process when it is used appropriate by teachers and is accompanied with meaningful feedback that helps students to reflect on their learning and make sense of their performance.

In other words, a good grade will be more significant once further explanations are provided to help the learner make connections between the grade and his or her strengths. This

is usually done during formative assessment where students are provided with information about their learning progress and suggestions for improvement (see Section 2.2.2 for formative assessment). According to Brown and Abeywickrama (2010), verbal and written descriptions can be useful in providing individualized feedback that informs individual students about their strengths and weaknesses and identify areas that need more attention and how they can be improved. Although teachers' measurement process is not of interest to this study, the context surrounding teachers' communication of measurement results will form part of current analysis focusing on the nature of the communication, its frequency and its intended purposes.

Unlike measurement that consists essentially of information quantification, the term *evaluation* is used to refer to the process of judging the obtained results for decision making. It is the process of determining the 'value' of test results (Bachman, 1990), and can take place after measurement or may happen without measurement. It is about making judgment and taking decisions on the basis of the obtained information and results (Airasian, 2005). In the evaluation process, the evaluator judges the learner's performance "in relation to a set of learner expectations or standards of Performance" (Cheng et al., 2004, p. 363). The educational evaluation process usually leads to decisions such as student promotion or presentation of awards. In other words, it is through evaluation that conclusions such as 'pass', 'fail' or 'reseat test' are taken.

2.1.2 Testing

In addition to *measurement* and *evaluation*, *testing* is another important term that is frequently used in education and in EFL contexts in particular. *Testing* is a part of the assessment process that involves the use of tests and a variety of testing procedures to collect information on the progress of the student towards the curriculum objectives (Airasian, 2005; Rea-Dickins and Germaine, 1992). A *test* in an educational setting refers to methods, the means through which data about the student's learning development is collected. Such methods need to be "explicit and structured" and should be able to provide the "accurate measure of the test-taker's ability within a particular domain" (Brown, 2004, p 4). Unlike most other means of assessment, tests are usually planned and structured in accordance with a given set of formal testing rules and administered according to the formalized procedures (Fulcher, 2012). The design of tests has

been changing particularly since the early years of the 20th century following the ever changing language testing paradigms. In the following section, a brief review of the testing paradigms is presented to allow for a full understanding of how testing as an institutional practice has evolved over the last decades.

2.1.2.1 Language testing paradigms

Institutional language testing began to be widely practiced in the mid-twentieth century when an important shift in testing approaches took place, moving from principle and procedure free testing period referred to as *traditional/pre-scientific* era (Spolsky, 1995) to the scientific testing era. According to Weir (1988), three main historical approaches in language testing are noteworthy: *Psychometric-structuralist era*, *Psycholinguistic-sociolinguistic era* and the *communicative approach*. As Ellis (2003) points out, testing in *psychometric-structural era* focused attention on reliability i.e. the objectivity and consistency of tests and the level of generalization of results. Testing was separated from teaching and no interest was put on assessing learners' language ability as a whole. Tests were centred on phonology, grammar and vocabulary all broken down in discrete elements that were measured in relation to speaking, writing, listening and reading. There was also a wide use of more controlled methods of assessment mainly true/false questions, multiple-choice items and word and sentence level translation from learners' native language to target language. With some researchers viewing teacher-made classroom tests as unreliable, the psychometric-structuralist type of tests gained momentum particularly after the publication of Robert Lado's work on language testing in 1960s which became influential for the thriving of standardized tests (e.g. the Michigan Test of English Language Proficiency in 1961).

However, the psychometric-structural approach to language testing was called into question in 1970s after criticisms mounted that the approach focused on discrete language elements and paid little attention to learners' communicative competence and authenticity of tasks (Brown and Abeywickrama, 2010). Although psychometric approach was credited for its potential to maximize test objectivity, it was criticized for failing to promote the core purpose of language teaching and learning which is to increase learners' ability to function well in real life communication. The *psycholinguistic-sociolinguistic era* of language testing also known as

scientific era (Madsen, 1983) thus emerged following following the decline of psychometric testing enthusiasm.

The *psycholinguistic-sociolinguistic* perspective viewed learners' language competence as a composite of linguistic skills that need to be assessed as a whole and in a more contextualized way. It argued that testing should be focused on real life language functions and as a result, be functionally and socially contextualized. In psycholinguistic sociolinguistic perspective, tests should be comprehensive in terms of language skills and at least two skills should be tested together to reflect the complexity of the real life communication (Oller, 1979). Within this context, the *integrative language testing* approach was introduced by linguists most notably Oller (1979) who stressed the need to test language as a wholeness. The approach argued for the use of holistic testing methods and authentic assessment to test the communicative competence and overall language proficiency. Cloze tests in form of dictation and gap filling text reading became popular under claims that the completion of these tasks required learners to use their overall language proficiency (Oller, 1979).

However, the *integrative approach* was criticized for focusing on learners' language competence and failing to test learners' language performance abilities (Weir, 1990. Carroll (1980) argues that task in integrative testing "does not represent genuine interactive communication and is therefore only an indirect index of potential proficiency in coping with day-to-day communication tasks." (p. 10). Some linguists also questioned the claim that cloze-tests which are common in integrative approach may suffice to measure learners overall language competency (Weir, 1988). This assertion was rejected on grounds that high dependence on cloze testing formats may instead affect some learners who would not be able to demonstrate their abilities through other testing techniques. In addition, there were concerns regarding the validity of dictation and cloze tests which do not offer opportunity to communicate spontaneously (Carroll, 1980). A new approach emphasizing the communicative purpose of language learning was thus introduced.

Viewed as an alternative to previous approaches, the *communicative language testing* advocated the validity of tests i.e. relevance of tests to language proficiency theories (Ellis, 2003) and use of authentic and practical assessment tasks (e.g Green and CCC, 1987; Weir,

1988). The major principle of this approach suggests that authentic and valid language assessment should not be hampered by the wish to meet reliability requirements. This is explained by Fulcher (2000) who maintains that communicative approach “was primarily a rejection of the role that reliability and validity had come to play in language testing.” (p.483). From a communicative perspective, testing was conceptualized as based on performance, authenticity and real life outcomes (Fulcher, 2000).

The increasing quest for valid assessment that measures learners’ performance in real-world like situations motivated the development of *performance-based* methods of assessment, also often called *task-based assessment*. The performance-based methods of assessment are still current and have been researched on in a number of studies investigating classroom language assessment (e.g. Bachman, 2002; Eckes, 2008; Mislevy et al., 2002). Proponents of performance-based assessment commend it for the high level of validity and authenticity of assessment tasks in which learners are engaged. According to Ellis (2003), validity of assessment is a correlation between what test takers do on tests (test performance) and what they have to do in real-life situations (criterion-performance). In other words, tasks in performance assessment are expected to be relevant, and mirroring real-life situations in which learners are likely to use their language skills outside the classroom.

The language testing approaches described in this section highlight the three major trends that saw language assessment emerge from the era of rule free testing to the scientific, objective testing period that focused on discrete linguistic knowledge and where reliability of tests was of high priority. The literature presented thus far provides evidence that language testing also evolved from this exclusive focus on test objectivity to learner centred assessment that puts emphasis on helping learners to develop their linguistic abilities and be able to communicate effectively in real life situations. Currently, language tests continue to occupy an important place in language teaching and learning. The following section explores the literature on the current place that the standardised and high-stakes tests continue to occupy in EFL educational contexts.

2.1.2.2 Use of language tests

According to Shohamy (1994), language testing can be situated in three main contexts: *SLA (second language acquisition) context*, *Classroom context* and *External context*. The *SLA context* is described as a context in which “language tests are used as tools for collecting language data in order to answer and test SLA research questions and hypotheses” (Shohamy, 1994, p.133). On the other hand, the *Classroom context* which consists of classroom-based assessment refers to a situation “where tests are used as part of teaching and learning process” (Shohamy, 1994, p. 133). These tests may be made by teachers (commonly known as teacher-made tests), adopted or adapted from other sources such as standardized tests usually designed for use in External or SLA contexts. Shohamy defines the *External context* as the context where standardized tests are used to evaluate individuals and programmes for purposes of certification, diploma awards, placement and acceptance or rejection to programmes. In this category are classified standardized or high stakes tests often developed by testing experts and usually designed for large-scale testing at school, state, or international level. The testing systems in many educational contexts are often put in place as a mechanism that aims to benefit learners and teachers and to inform other educational stakeholders in order to better serve the national interest as a whole. However, the use of tests can serve many other purposes.

According to Supovitz (2009), attachment to assessment in general and high stake testing in particular is often linked to motivational, informational, symbolic and alignment theories that explain the use of high-stakes tests in educational context:

Four major theories underlie our current reliance on high-stakes tests: *motivational theory*, which argues that test-based accountability can motivate improvement; *the theory of alignment*, which contends that test-based accountability can spur alignment of major components of the educational system; *information theory*, holding that such systems provide information that can be used to guide improvement; and *symbolism*, which maintains that such a system signals important values to stakeholders. (p.2)

Consistent with Supovitz (2009) claim, testing in some contexts is designed, at least theoretically, to provide information on the effectiveness of educational programmes for further decision taking. This is the case for Rwanda where English language assessment is done through the national high stakes test officially for purposes matching the four testing reasons highlighted above. The national examination administered at the end of the F3 of Rwandan

secondary education level, which is the scope focus of this study, is one of two major national tests that are taken by secondary school students for promotional and certification purposes among others. According to MINEDUC (2010), this exam is designed to yield eligibility for the third year students to be promoted into the fourth year and give a comprehensive picture of how well students are successful in learning.

However, this is a high-stakes test that has political and administrative implications. For example, teachers and head teachers sign the annual performance contract in which a pledge is made to increase the number of students performing well on the test. In some cases, teachers and head teachers of some schools have been obliged to resign as a result of failing to meet their contract performance objectives including mainly boosting students' performance on national tests (Kamanzi, 2012). In fact, beyond the intuitive measurement and evaluative purposes, high stakes testing systems play other significant roles that also shape the way teachers conduct their classroom assessment. This is emphasized by Shohamy who argues that tests are often used as political and administrative instruments. She states that "Policy-makers in central agencies, aware of the power of tests, use them [tests] to manipulate educational systems, to control curricula and to impose new textbooks and new teaching methods" (Shohamy, 1993, P. 5).

Although most assessment systems and frameworks used in schools are officially designed for measuring language learners' progress and as a leverage for developing language learners' communicative competence necessary for international market competitiveness (see Ahn, 2015), some policies also often carry hidden managerial and policy administration agenda (McNamara, 2001) and can be as fundamental as defining knowledge focus of the classroom instruction. Shohamy (1997) describes the purpose of tests use as multidimensional, emphasizing that being embedded in educational, political and social contexts, tests particularly standardized tests are powerful tools that are used to "define and impose language knowledge and create de facto language policies" (p. 522). In her study, Hsu's (2009) interviews with the administrators at a Taiwanese university revealed that the examination introduced as a graduation requirement was set up to ensure that English was taught at the university. Broadfoot and Pollard (as cited in Rea-Dickins, 2007) argues that interest in assessment in some countries has been motivated by a number of factors including mainly economic

competition, financial pressures and an increase in demand for accountability. In the following section, further review of the influence of language testing policies is presented. This analysis allows for advanced understanding of the extent to which both the high stakes tests and the accompanying hidden agendas affect the EFL teacher's classroom assessment practices.

2.1.3 Assessment

Although *testing* and *assessment* are sometimes used interchangeably, the term *assessment* covers a broad range of parameters than *testing* and it is used usually as an umbrella term for *testing* and *measurement* (e.g. Coniam and Falvey, 2007; Rea-Dickins and Germaine, 1992). As briefly stated earlier in the introduction to this chapter, assessment in educational settings is defined on the whole as the process of eliciting information about the learning progress of students. This process can use tests or other methods of information gathering such as observation, interview and behaviour monitoring (Overton, 2012). In this study, the focus is on assessment taking place in classroom often referred to as *classroom-based assessment* or simply *classroom assessment*. This is usually prepared and administered by teachers as an integral part of their everyday instructional practice.

2.1.3.1 Classroom assessment

Classroom assessment refers to the continuous classroom process of collecting facts about the learning development of students in a given instructional setting. However, this definition may appear too simplistic and does not reflect the existing different interpretations of the term 'assessment'. These interpretation differences of classroom assessment are, as Rea-Dickins (2007) explains, influenced by different interpretations of the purpose of assessment and practices associated with it. For example, Brown, Hudson and Clapham (as cited in Rea-Dickins 2007) define classroom assessment in terms of formality and reliability referring to the use of assessment standards and formal procedures. Such standards and formal procedures are often characteristic of high stakes testing viewed by some educators as "detached from real learning and real-life performances" (Shohamy, 2005, p. 103).

On the other hand, a differing perspective views classroom assessment as a less formal process "where learner performance is analysed in terms of learning goals and instructional

process rather than a finished product” (Rea-Dickins, 2007, p. 507). This refers to formative assessment (see Section 2.2.2) which is part of the teaching and learning process and aimed at improving both learning and teaching. In view of the current literature, the classroom-based assessment can be seen as complex and hard to find a definition that reflects all its forms. This complexity is reflected in what Rea-Dickins (2001) call “assessment opportunity” defined by Hill and McNamara (2012) as “any actions, interactions or artifacts (planned or unplanned, deliberate or unconscious, explicit or embedded) which have the *potential* to provide information on the qualities of a learner’s (or group of learners’) performance.” (p.397). Unlike the other definitions centred on well-known formal and informal practices, the definition by Hill and McNamara also takes into account the classroom assessment forms that are used instinctively and therefore often ‘invisible’ taking place in ‘real time’ in the classroom day to day interaction.

In this study, classroom assessment is used to refer to the definition by Hill and McNamara (2012) that classroom assessment is “any reflection by teachers (and/or learners) on the qualities of a learner’s (or group of learners’) work and the use of that information by teachers (and/or learners) for teaching, learning (feedback), reporting, management or socialization purposes.” (p. 396. This definition is maintained in this study on the basis that it can fit within different assessment contexts as it encompasses a wide range of assessment purposes and takes into account both the use and non use of assessment results. The following sections re-examine the existing literature on the classroom assessment practices and explore the latest research findings on how teachers decide on what they do in the classroom. Research findings suggest that both teacher-made and standardized tests are some of the key forms of assessment that teachers use for different purposes in different classroom assessment settings (Dunn et al., 2003; Tsagari and Pavlou, 2009). As briefly discussed above in section 2.1.2.2, the use of standardised tests in the classroom setting has some major implications. The following section explores these implications and reviews literature on the influences that the use standardised and high stakes tests have on the practices of assessment at the classroom level.

2.1.3.2 Impact of high-stakes tests on classroom assessment

Over the last two decades there has been a growing disenchantment against high reliance on standardized testing and critics have emerged against assessment systems that are incompatible

with instructional contexts. It has been argued that the use of high stakes tests can be a major obstacle to the effective classroom assessment practices (Rea-Dickins, 2008) and some educators have advocated for institutional support to the teachers to help them resist the influence of standardised tests in their everyday classroom practices (Shepherd, 2000). Some researchers have cautioned against testing systems in some contexts where such systems only endorse other agendas and do little to promote effective teaching and learning (e.g Shohamy, 1997):

Policy makers in central agencies use tests [...] to manipulate educational systems, to control curricula, to create new knowledge, to define knowledge and to impose new textbooks, to communicate educational agendas and new teaching methods. Principals use school-wide exams to drive teachers to teach and teachers use tests and quizzes to motivate students to learn and to impose discipline. (p. 347).

Shohamy (1997) also considers the use of high stakes tests as unethical in situations where teachers are compelled to rely on such tests for instructional content:

In a high stakes situation teachers react by teaching the topic. Since teachers have no knowledge of the new topic they turn to the most immediate pedagogical source, the test itself, to learn how to carry out these new demands. The test thus becomes the new de facto knowledge and thereby the device through which control is exercised, legitimizing the power of bureaucrats and elite groups. It is my claim that the use of tests in these ways is unethical. (p. 347)

There is currently a considerable amount of literature showing that significant consequences both positive and negative, wanted or unwanted result from the use of inflexible assessment policies and overreliance on summative testing systems (Azadi and Gholanmi, 2013; Cheng and Watanabe, 2004; Cheng, 1996; Munoz and Alvarez, 2010; Nkosana, 2008). These studies came to be known as washback studies that investigate the impact of assessment in general and of high stake tests in particular. Over the last few decades, the washback effects of standardised tests have been widely researched.

Studies on the impact of assessment started to assume an important place in the educational research realms in the 1990s with the publication of works on tests washback effects (see Alderson and Wall, 1993). Although the term washback was already in use in educational research, it is after Alderson and Wall's study investigating the existence of washback that a research agenda on test washback was initiated. Originally also called

'backwash', washback is defined as the impact that tests have on teaching and learning with a common assumption that good tests will have good impact (Alderson and Banerjee, 2001). In more than a decade now, studies have explored the test washback phenomenon with a particular focus on high stakes testing (Harlen and Crick, 2002; Hsu, 2009; Watanabe, 2004). It has been repeatedly found that washback is a complex phenomenon whose intensity and nature are highly context dependent. Although most studies have focused on effects of tests on teaching practices or the behavioral, attitudinal and educational aspects (e.g. Ryumon, 2007; Sara and RuiLei, 2013; Shohamy, Donitsa-Schmidt and Ferman, 1996), some researchers have also looked at tests impact on classroom assessment with findings showing that standardized/high stakes testing affects assessment and testing practices taking place in classrooms. Hsu (2009) used a questionnaire survey, interviews and classroom observations to investigate the impact of introducing an English test as a graduation requirement at one Taiwanese university and found that the introduction of an English test requirement resulted in significant changes in teachers' classroom assessment practices. Teachers started to use cram tests reflecting content items on the graduation tests in addition to their own developed tests. In addition to Hsu's findings, similar conclusions were made in a study by Shim (2009) who found that one in five teachers of English in Korean primary schools used external assessment materials including those from standardized tests.

In addition to studies that have focused specifically on washback, a number of research studies that have investigated the language testing and/or classroom assessment have also touched upon issues of the effects of large scale testing on classroom assessment. It has been found that the prevalent use of external high stakes tests affects classroom assessment practices with direct effect on learners (Bailey, 1999; Read and Hayes, 2003; Shohamy et al., 1996). Studies have indicated that being summative in nature and serving normative purposes, standardized testing can have deleterious consequences and cause serious emotional and affective toll on learners. Low performing testees may be left with a feeling of incapability while having to compete with highly performing colleagues. Moreover, because of the high stakes attached to summative tests, teachers are usually inclined to adopting and/or adapting the norms and procedures used in such high stakes tests irrespective of possible inappropriateness. Black and Wiliam (1998) argue that by adopting the methods of standardized tests, some

teachers give up on their power and flexibility that normally help them tailor classroom assessment tasks to the needs of students:

Such tests can dominate teachers' work, and insofar as they encourage drilling to produce right answers to short out-of-context questions, this dominance can draw teachers away from the paths to effective formative work. They can thereby constrain teachers to act against their own better judgement about the best ways to develop the learning of their pupils. (p.147)

As was stated earlier in this chapter, most high stakes tests are usually designed to meet the traditional criteria of reliability and validity which, as discussed in earlier sections, present considerable weaknesses and are often not compatible with suggested formative assessment practices. For example, it is unsurprising that some classroom assessment practices that were widely used during the structural linguistics era (such as the extensive use of multiple-choice) can still be observed in current classroom assessment practices even in situations that seem less appropriate (see Cheng et al., 2004; Shim, 2009). These practices are usually frequent in contexts where high stakes national or local level tests are used solely or as part of a wider assessment framework. In the context of this study, issues of high stake testing influence on classroom assessment are relevant and worth examining for a thorough understanding of factors influencing classroom assessment practices. Teachers' classroom assessment practices in Rwanda, a country that relies on examinations to measure the effectiveness of programmes (USAID, 2010), are likely to be influenced by high stakes national exams. The existence of one large scale national test at the end of the third year of lower secondary school bears a great importance for teachers, schools and learners and it is against the same test that the performance of teachers and head teachers in lower level secondary schools is measured (Kamanzi, 2012). This underscores the warrant to explore assessment practices that teachers in second and third years of this educational level employ in classroom and this study seeks to understand further how such practices impact on the students' learning motivation especially with regard to their self beliefs.

The literature presented in this section demonstrates that high stakes, large scale tests can have some negative effects on the classroom-based assessment. However, it has also been argued that these tests can have positive impact on the classroom assessment and can be used to advance learning through formative assessment as Black and Wiliam (1998) stress:

This is not to argue that all such tests are unhelpful. Indeed they have an important role to play in securing public confidence in the accountability of schools. [...] what is needed is that in the evaluation of such tests, and in any development programme for formative assessment, the interactions between the two be studied with care to see how the models of assessment that external tests can provide could be made more helpful. (p. 147)

Although it also appears that high stakes tests can play significant role in advancing learning at the classroom level, they seem to remain effective for serving the high level agenda and are far from being a primary model for teachers in their everyday classroom assessment. As a result, researchers and practitioners have advocated some practices that are seen as effective in supporting learning. In the following section, recent literature on the effective practices of classroom assessment is reviewed.

2.2 The case for effective classroom assessment practices

Early advocates of classroom-based assessment became prominent in the late 1990s particularly after the publication of a review on classroom assessment by Black and Wiliam (1998). In their review, Black and Wiliam concluded that teachers' effective practices of classroom assessment can help raise the standards of achievement. Black and Wiliam (1998) pointed out that such practices should be aimed at contributing to the improvement of classroom assessment, hence reemphasizing the concept of formative assessment, also often known as 'assessment for learning' which places teachers classroom assessment at the centre of classroom activities. The classroom assessment was expected to shift from the traditional outcome-based assessment to process-based assessment focused on supporting the learning progress of students.

In Figure 2.1, McMillan (1997) draws our attention to a number of other assessment trends that have emerged since 1990s and which have emphasized the importance of putting learners at the centre of assessment practices. Assessment practices associated with these trends have been encouraged over the past decades with a particular focus on assessment for learning practices or *formative assessment* (see Section 2.2.2). Many of these new trends are also characteristic of 'Assessment for Learning, which emerged as a response to the overwhelming

role that assessment policies and external testing were gradually assuming in the field of assessment.

Trends in Classroom Assessment	
From	To
Sole emphasis on outcomes	Assessing process
Isolated skills	Integrated skills
Isolated facts	Application of knowledge
Paper-and-pencil tasks	Authentic tasks
Decontextualized tasks	Contextualized tasks
A single correct answer	Many correct answers
Secret standards.	Public standards
Secret criteria	Public criteria
Individuals	Groups
After instruction	During instruction
Little feedback	Considerable feedback
"Objective" tests	Performance-based tests
Standardized tests	Informal tests
External evaluation	Student self-evaluation
Single assessments	Multiple assessments
Sporadic	Continual
Conclusive	Recursive

Figure 2.1: Recent trends in classroom assessment (from McMillan, 1997)

The interest in learner-centred assessment practices increased among educators and researchers who developed a new view of the purpose of assessment in general and of classroom-based assessment in particular. In contrast to the use of traditional psychometric methods that were often used to measure learning achievement, a remarkable interest was directed towards assessment that not only aims to report learners' achievement but also has as the primary purpose to improve learning. This led to the embrace of classroom assessment that uses 'authentic and contextualized tasks' and gives 'considerable feedback' to learners (McMillan, 1997).

Being learner centred, the classroom-based assessment has been thus viewed by a growing number of researchers and educators as an effective alternative to standardised testing thanks to its aims for instructional and learning improvement, learner centeredness, focus on individual learner and potential of learning opportunities and boosting the learning motivation of learners. Over the years, assessment practices that improve both the instruction and learning have been viewed as characteristic of effective classroom assessment (Black and Wiliam,

1998). Typical of effective classroom assessment are also constructive feedback that inform students on their strengths and weakness and that suggest ways to improvement (Hattie and Timperley, 2007). In addition, effective classroom assessment allows for the use of a wide range of assessment practices that fit every learner's way of demonstrating their learning progress and reveals their weaknesses and strengths. In short, the practice of classroom assessment can be carried out in many different ways in order to maximize the learning potential of every learner. Despite the multidimensional aspect of the classroom assessment practices, some frameworks have been suggested as a way of standardizing the practice. In the following sections, the discussion focuses on one framework by Hill (2017) emphasizing the importance of carrying out assessment in more structured way. The three-element framework represents the knowledge and skills that teachers need for their classroom assessment. The following sections review some key issues that have dominated the discussion pertaining to these major elements of the classroom assessment practices. This review highlights the current theoretical underpinnings in support of learner-centred classroom assessment and the focus is particularly put on the current literature on the case for the use of diversified forms of assessment, the importance of formative assessment, the role of feedback in promoting learning.

2.2.1 A three-element framework

The discussion of assessment practices tends to be focused on the standardised testing that use large-scale and usually high-stakes tests. However, it is important to reiterate that by using the term 'classroom-based assessment' the emphasis is on assessment carried out "by teachers (and/or learners) on the qualities of a learner's (or group of learners') work and the use of that information by teachers (and/or learners) for teaching, learning (feedback), reporting, management or socialization purposes" (Hill and McNamara, 2012, p. 396). The classroom assessment practices may differ from teacher to teacher and from subject to subject. However, some practices have been recommended for effective classroom assessment in educational settings. As Hill and McNamara point out, the practice of classroom assessment is a multifaceted process that involves the execution of tasks by learners, at the completion of which relevant information is obtained and can be used by learners and teachers. This practice of classroom-based assessment is further explained by Hill (2017) as he presents the different

skills that teachers need to acquire for the development of their assessment literacy. The framework is based on three dimensions namely *practice* i.e. what teachers do for the classroom-based assessment, *concepts* i.e. in relation to the theories and standards that guide the teachers' assessment and *Context* i.e. regarding the ways in which the teachers' assessment practices are influenced by the context of their work. In the current study, the focus is on the practice, the different ways in which teachers accomplish their assessment roles in their respective classrooms.

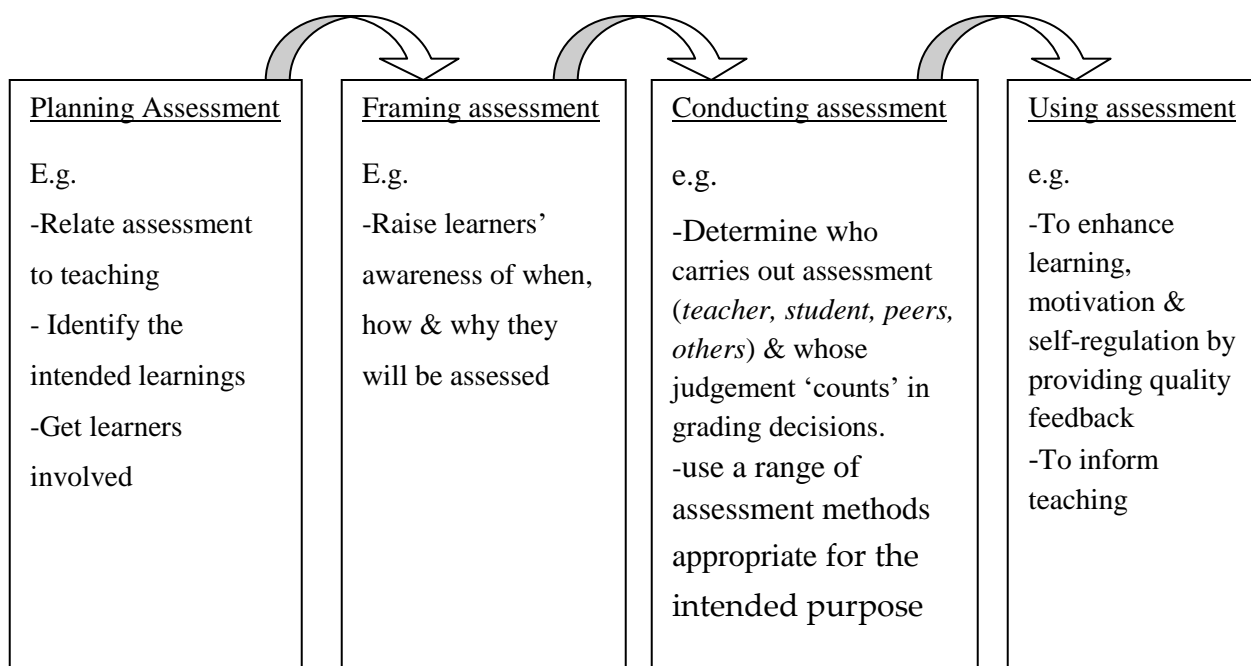


Figure 2.2 Elements of classroom based assessment practices (Adapted from Hill, 2017, Pp. 5-6)

As indicated in Figure 2.2, Hill (2017) identifies three elements characteristic of the assessment practices in the language classroom. The first element is the *planning of assessment* where teachers examine whether their assessment is related to the learning targets and the extent to which the learners' related factors are considered. These may include the learners pre-existing language knowledge and their language needs towards which learning objectives are oriented. Learners also need to play their role in planning assessment to develop their interest and learning ownership (Title, 1994). For Anderson (2003), the planning stage is also the time when teachers decide on the timing and the purpose of the assessment. The *Framing* element is concerned with the ways that teachers use to inform the learners about assessment. It may consist of communicating the venue, timing and assessment objectives of the assessment. Central to the classroom assessment practice is the way assessment is carried out (*conducting*

assessment). In this phase, Hill underlines the importance of identifying and effectively using appropriate methods for assessment in order to meet the assessment objectives. Clarification on the assessment participants and their role is also essential and measures to ensure quality of assessment need to be put in place. The fourth element of the classroom assessment is *using assessment*. According to Hill (2017), this consists of using the information obtained at the end of assessment tasks for different purposes such as measuring the learners' progress toward the attainment of learning objectives, reporting to stakeholders or helping the students to prepare for exams.

2.2.2 The formative purpose of classroom-based assessment

Formative assessment is defined as assessment “encompassing all those activities undertaken by teachers, and/or by their students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged” (Black and Wiliam, 1998, p.7-8). The modification of teaching is done to respond to the identified learning needs of students and to help them achieve their learning goals. Learning and assessment activities take place during the course time and they do not have to be formal. Unlike summative assessment, formative assessment uses more informal methods to gather information on students' performance (Harris, 2005). The formative assessment is often contrasted with summative assessment which is also referred to as ‘assessment of learning’ (Earl, 2010; WNCP, 2006). This usually takes place at the end of a course unit or at the end of a term to measure the learning achievement of students (Harris, 2005). In summative assessment, formal methods are used to collect evidence of students' attainment and grades or ranks are often assigned to the learners' work to add value to their performance for important decision making (Alsop and Ryan, 1996).

In addition, while summative assessment is often believed to be focused on measuring the amount of what is learned and serves as the basis for final educational decision making, formative assessment is credited for being centred on feedback which is considered as fundamental to the teaching and learning process (see Section 2.2.2). It is important to note that standardized tests are usually summative. Taking place at the end of an instructional programme or at the end of school year, Summative tests are often used for important academic

decision making. It is worth noting, however, that the distinction between formative and summative assessment is not always explicit since summative assessment can also sometimes be used for formative purposes. In their study on the nature of formative assessment in language learning in UK elementary school, Rea-Dickens and Gardener (2000) observed that, in some cases, assessment results obtained at the end of the term were used for instructional planning in the next term.

Given its emphasis on goal-oriented classroom assessment activities carried out to support immediate learning, formative assessment is sometimes used interchangeably with classroom assessment or ‘assessment for learning’. Black and Wiliam (1998b) indicate that for classroom assessment to be formative, the result from assessment activities must be used to bring about improvement in learning and teaching. This suggested relationship between assessment and learning is at the origin of the term ‘assessment for learning’ which focuses on the process (learning) rather than the product (achievement) (Stiggins, 2002; Earl, 2010). This process is usually based on interactions both between teacher and students and among students themselves. In formative assessment environment, learning intentions are clarified and shared with learners who take ownership and become learning resources for one another (Black et al. 2004), supported by teachers who provide feedback geared towards moving learners forward, in order for them to achieve intended curricular aims (Popham, 2008; Wiliam, 2011). The use of formative assessment helps identify areas in which learners experience learning problems and assist teachers to reset the teaching targets and learning outcomes where students are informed, through effective feedback, about their strengths and shown where more efforts are needed.

2.2.3 The role of feedback in classroom-based assessment

In their everyday classroom assessment, teachers provide feedback on learners’ performance to encourage them to reflect about their work and suggest different ways for improvement (Rust, 2002). Being an essential part of any educational assessment process, feedback is centred on the generation of information about the progress of learning. As shown in Figure 2.3, the purpose of feedback is to help narrow the gap between the learners’ actual understanding or performance and the targeted learning objectives (Hattie and Timperley, 2007). Teachers

collect information on an everyday basis on what goes on well with the course and what needs immediate modification for the purpose of making effective learning happen. Hattie and Timperley highlight in Figure 2.3 how the collected information can be used by both the students and the teachers to fill the gap between understanding and learning goals. This is done when assessment information is used by teachers and students to adopt effective strategies that are suitable to the learners' needs. The moment of feedback provides an opportunity for teachers to help learners know where they are in the learning process and help them build the necessary confidence and get more encouragement in learning activities.

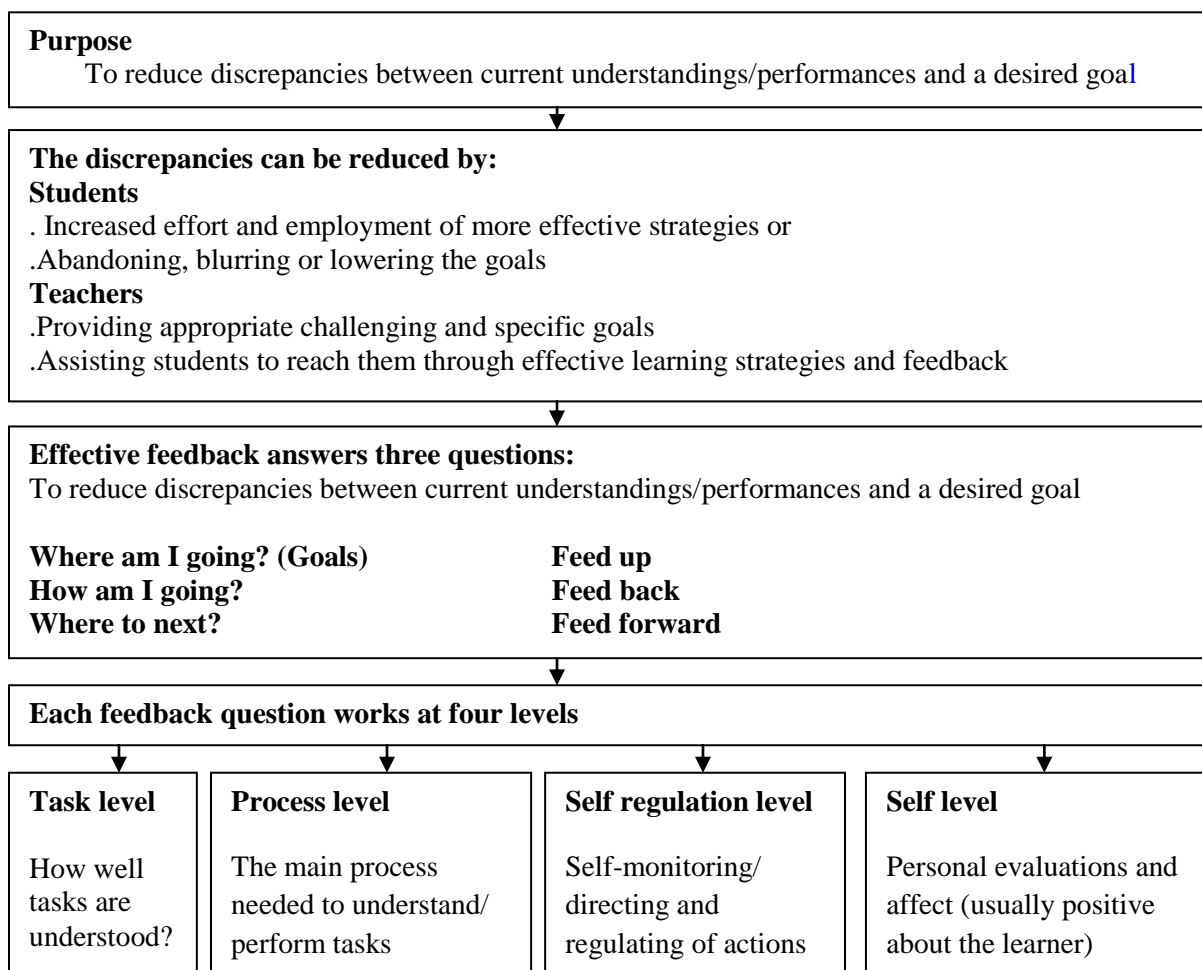


Figure 2.3: A model of feedback to enhance learning (from Hattie and Timperly, 2007)

For more than two decades, more evidence has indicated that promoting and strengthening assessment feedback leads to significant achievement (Black and Wiliam, 1998, 2001; Hattie and Timperley, 2007). Research on language learning and assessment has

indicated that effective feedback is the central part of any formative assessment that contributes to both learning and teaching. After Black and Wiliam's (1998) review, other studies have provided more empirical evidence of the effectiveness of formative assessment and this has attracted attention of educators and policy developers who have started to incorporate formative assessment in assessment policies (Leung and Rea-Dickens, 2007). It is also argued that frequency of formative assessment activities correlate with the learners' achievement and that the more frequently effective assessment incorporates timely feedback, the more likely that effective learning will take place (Marzano, 2006).

2.2.4 Forms of classroom-based assessment

In their everyday classroom assessment practices, teachers use a variety of assessment means many of which engage learners in different tasks at different times. In classroom assessment research, different terms have been used to refer to ways by which teachers collect evidence of learner progress in the classroom. Terms such as '*methods*' (e.g. Butler and McMunn, 2006; Stiggins, 1991), *tools* and *instrument* (e.g. Brookhart, 2010; Lorin, 2003) or assessment *techniques* (Phye, 1996), are sometimes used interchangeably. In this study, the term "*form of assessment*" is used to refer to different means by which teachers obtain information about their learners' progress in the classroom. These may be in the form of quizzes, written tests, oral questioning, tests, interview, debates or oral presentations among others. The term "*form*" is chosen to avoid the notion of 'procedure' in the term '*method*' or the idea of '*device*' that can be understood in terms 'tools' and 'instrument'. The term *form* reinforces this study's emphasis on the ways of assessment, structure, the configuration or shape in which assessment appears in the classroom.

There is a wide range of assessment forms that teachers use to offer opportunities to learners to demonstrate their capacity and to get the chance to experiencing success. Unlike in the case of standardised tests where well established psychometric assessment methods are often used on a large scale, classroom assessment allows teachers to decide on the appropriate form of assessment to use for eliciting reliable information on students' learning progress. Assessment can be carried out in different forms depending on the appropriateness to the learning context and the suitability for gathering accurate information that can help teachers

adjust their instruction and students to fine-tune their learning tactics. For many decades, researchers and language educators have investigated the forms used by teachers to assess their learners and a growing body of literature has offered significant insights into the many ways that teachers use to collect information from learners (e.g. Cheng et al., 2004; Hill and McNamara, 2012; Inbar-Lourie and Donitsa-Schmidt, 2009).

The use of multiple forms of assessment has been considered essential in promoting learning as it enables teachers to “triangulate” evidence and to have a complete picture on the learners’ progress towards the learning objectives (Butler and McMunn, 2006). According to Butler and McMunn, learner progress information can be collected in two forms: (1) by using general ways that include teacher observation, questioning and teacher-learner dialogue or (2) by using a variety of specific ways such as multiple choice, essays or oral presentation. The specific ways of assessment listed by Butler and McMunn are usually classified in two categories that may take different names depending on the preference of researchers. In some studies, the term ‘*traditional assessment*’, [probably deriving from its long history of use compared to other forms] and ‘*alternative assessment*’ are used to refer to those forms of assessment comprising closed question items and open ended items respectively.

In traditional assessment, learners are presented with response options from which to choose the right answers (‘*selected responses*’ according to Butler and McMunn). These include question items in the form of true/false, matching, fill in or multiple choice questions (Rhodes, Rosenbusch and Thompson, 1996). The term *paper-and-pencil* assessment is also often used to refer to *traditional assessment* (Stiggins, 1991). On the other hand, the term *Performance-based* assessment is used to refer to assessment forms such as oral presentation in speaking or some composition tasks such as poem writing (Rhodes et al., 1996). The term ‘*alternative assessment*’ is also often used to refer to open ended type of question items that requires learners to demonstrate knowledge actively through tasks that often involve performance. The *traditional paper and pencil* based forms of assessment are often associated with assessment of linguistic skills. They are commonly used in standardised tests as they are believed to increase objectivity and fair scoring.

However, some researchers have argued against reliance on paper and pencil assessment claiming that they fail to “address important curriculum goals that require generative thinking, sustained effort over time, and effective collaboration” (Ananda and Rabinowitz, 2000, p. 2). In contrast, performance –based assessment is thought to be appropriate for assessing the skills as Ananda and Rabinowitz (2000) further point out: “ These important skills are better assessed by performance-based assessment methods, such as portfolios, computer simulations, oral presentations and projects, which make greater demands on academic foundation, teamwork, and problem-solving skills than traditional paper-and-pencil assessments.” (p. 3). The use of performance-based assessment has become particularly popular following the increasing interest of educational practitioners in communicative language teaching (see Brown and Abeywickrama, 2010).

In the current study, the term *paper-and-pencil* based assessment is used to refer to forms of assessment comprising closed, more controlled question items usually done on paper. These consist of objective assessment where learners have to “recognize rather than generate answers or create brief responses to questions in which they have little personal investment” (Windschitl, 1999, p. 753). On the other hand, ‘*performance assessment*’ is used in the current study to refer to forms of assessment such as oral presentation, role plays and debate. These *performance assessment* forms provide alternatives to learners to demonstrate their knowledge in many different ways. Stiggins (1991) clarifies that teacher made tests, quizzes, homework and seatwork exercises and assignments can all be used for *paper-and-pencil* assessment to measure the learners’ attainment. Tests may also be sourced from textbooks and published standardized tests. For *Performance* assessment, Stiggins gives examples of observations of the learners’ behaviors and the evaluation of individual learner’s product.

2.2.5 The use of Rubrics in classroom assessment

Essential to quality classroom assessment and particularly relevant to performance assessment is the use of rubrics. Rubrics are often referred to as documents essentially used for collecting information on the students’ progress toward specific learning targets and to enhance the learners’ performance (Arter, 2000). They are particularly considered as indispensable for performance assessment where teachers need to set up clear criteria upon which they judge the

quality of the students' performance (McTighe and Ferrara, 1998). The relevance of rubrics to performance assessment is highlighted in most literature (e.g Andrade, 2000; Mabry, 1999; Outeiral, 2014). As Andrade (2000) stresses, rubrics are used "to provide students with more informative feedback about their strengths and areas in need of improvement than traditional forms of assessment do." (p. 15). By using the traditional assessment methods often in the form of paper and pencil based assessment; teachers may not feel a need to use rubrics as answers to the question items are usually single and short. This is in contrast with performance assessment where well specified criteria are used to judge the quality of the students' work (Arter and McTighe, 2001).

Brookhart (2013) identifies two levels at which rubrics can be classified. First, rubrics can be classified as 'Analytic' when they are designed to deal with assessment criteria one by one separately, or 'holistic' depending on whether they are used to deal with all the criteria together. They can also be classified as 'general' when they can be used for several similar tasks, or 'specific' in the case they are used for one assessment task. Although each of these types of rubrics can have its advantages and disadvantages, Brookhart suggests that analytic rubrics are more appropriate for the classroom assessment as they guide the students on the specific aspect of their work that needs their attention. Rubrics can help reduce the students' learning anxiety (Andrade and Du, 2005) and can raise the students' self-efficacy by clarifying the target learning goals and assessment standards (Schunk and Zimmerman, 2007).

While the use of rubrics in educational settings has been increasingly recommended, it has also been argued that the use of rubrics has its own pitfalls and poorly prepared rubrics can be detrimental to learning. Cooper and Gargan (2009) summarises the disadvantages of using rubrics highlighting three issues that may rise from inappropriate use of rubrics. They argue that rubrics can be subjective when teachers use them "to convert lists of qualitative terms, each critical and in- dependent, into a set of scores that can be summed, averaged, and transformed into a grade" (Cooper and Gargan, 2009). They also point out that designing good rubrics can be time consuming for teachers who may already be overloaded with their teaching responsibilities and that the limited time for careful preparation may lead to poorly designed rubrics that restrict understanding and inhibit learning:

Both poorly designed rubrics and highly prescriptive ones can stifle students' creativity as teachers measure student work strictly by the rubric. Rubrics can become the overbearing framework that shapes student work, forcing everyone to look at problems and solutions in the same way, thus discouraging new ideas and approaches. Creative students - those thinking "outside the box" and beyond the rubric - will be penalized. (p.55)

However, the significant role that the appropriate use of rubrics plays in helping learners to understand and participate in the assessment process seem to outweigh these drawbacks. In addition, different ways have been suggested for the design of effective rubrics. Brookhart (2013) suggests that teachers who want to design effective rubrics need to think about and define clearly the learning targets that the students should aim to achieve. She argues that answers to the following two main questions can guide teachers in their process of clarifying the learning targets and describing them on the rubrics: "What are the criteria for good work on the task that the rubric is to assess? What should a student, peer, or teacher be looking for?" (Brookhart, 2013, p. 23). Andrade (2000) comments that instructional rubrics need to have two features: (1) Criteria or a list of what counts in a project or assignment and (2) quality levels of the student work indicating strong, middling, and problematic work.

2.3 Recent empirical research findings on classroom assessment practices

The recent years have seen a gradual shift from the historical research focus on standardized testing that focused on tests construction - reliability and validity- to learner centred research examining classroom assessment practices and their impact in various language teaching contexts. This research trend has resulted in a renewed interest to investigate issues of classroom assessment from socio-cultural, theoretical and practical perspectives. At the heart of recent classroom assessment research are issues of assessment processes (Davison 2004, Hill and McNamara, 2012; Rea-Dickins, 2001), the epistemological basis, practices and role of formative assessment and diagnostic classroom assessment (Alderson 2005; Black and Wiliam, 2008; Leung and Mohan, 2004; Lantolf and Poehner, 2010; Read, 2008), grading practices and decision taking (Brinley, 2001; Cheng and Wang, 2007), Classroom assessment methods (Brown and Hudson, 1998; Cheng et al., 2004), teachers' conceptualization of assessment (Cumming, 2001; Shim, 2009; Inbar-Lourie and Donitsa-Schmidt, 2009), and teachers and

student role in classroom assessment (Seong, 2011). In these studies, classroom assessment has been generally described as a complex process carried out for different purposes with practices heavily dependent on teachers' beliefs, classroom environment, and educational policies among other contextual factors.

Results from relevant empirical research have shed more light on the complexity of classroom assessment and demonstrated a multitude of factors that exert influence on the teachers' assessment practices. Some of the identified influencing factors include educational training and perceived competence (Alkharusi, Aldhafri and Alkabani, 2014; Chen and Sun, 2015; Yang, 2008; Zhang and Burry-Stock, 2003), ideological and political constraints (Inbar-Lourie and Donitsa-Schmidt, 2009) and teachers' limited English proficiency (Cheng et al., 2008). In their questionnaire survey of 113 elementary and high school EFL teachers in Israel, Inbar-Lourie and Donitsa-Schmidt (2009) found that technological factors were the most influential to the teachers' perceptions and usage of alternative assessment in their classroom. They also found that assessment complexity reflected the technological and political micro constraints as well as ideologies and commonly-held beliefs of teachers. In another correlational study, Yang (2008) investigated the factors affecting the use of multiple classroom assessment practices of 141 elementary EFL teachers in China. The findings indicated that teachers' assessment practices positively correlated with their perceived assessment competency, their perceived pedagogical benefits of assessment and their previous education. Their practices also negatively correlated with their perceived difficulty of implementing assessment.

Several research studies have also used different research methods to investigate the teachers' practices of classroom assessment in different contexts (e.g Frey and Schmitt, 2010; Öz, 2014; Tante, 2010; Tsagari and Pavlou, 2009). In their investigation of the classroom assessment practices of 191 EFL teachers from primary, vocational and high schools in Cyprus, Tsagari and Pavlou (2009) found that teachers reportedly assessed for different purposes and heavily relied more on paper and pencil format than alternative assessment. Their questionnaire survey showed that the teachers used limited varieties of assessment methods and assessed limited language skills. Teachers favoured external examination and textbooks and not shared assessment among themselves. Although the teachers reported using a variety of feedback, data

also showed that the focus of their feedback was on students' linguistic knowledge and emphasised the language 'product' rather than 'process'. These findings were similar to the results from Frey and Schmitt (2010) who, in their survey investigation of 140 3rd to 12th grade teachers in Kansas USA, found that the teachers rarely used formative assessment. The results of their study also indicated that more paper-and-pencil assessment forms were more common than performance based assessment.

In view of these findings and other results from recent studies, it can be observed that with regard to the forms of classroom assessment, teachers still rely more on paper and pencil forms than on performance based forms (Tsagari, and Pavlou, 2009; Zhang and Burry-Stock, 2003) and assessment for summative purposes (e.g Frey and Schmitt, 2010; Öz, 2014). Overall, the current literature highlights the important role that factors such as teacher beliefs, school environment and the social and political constraints in shaping the teachers' assessment practices.

2.4 Overcoming the challenges in classroom-based assessment

Recent theoretical works and research literature have made significant contribution in increasing our knowledge of teachers' assessment practices. However, there are still uncertainties and unanswered questions on a number of theoretical, political, social and most importantly practical issues regarding classroom assessment in general and in EFL I particular. There is also still a need to establish clear boundaries between assessment activities that constitute real assessment for learning and assessment activities that serve other purposes. The usual gap between national teaching agendas and assessment policies and the requirements of formative assessment is also an issue of high importance. This adds to the technical and practical challenges that arise during actual classroom practices of formative assessment relating to time management and lack of sufficient interactional skills not to mention the insufficiency of the material resources. In the following sections, a brief review of the existing literature on some of the key issues relating to the current practices of classroom-based assessment is presented. Particular attention is focused on issues relating to the current study

specifically the teachers' assessment literacy (Fulcher, 2012; Hill, 2017; Newfields, 2006), Issues of validity and reliability of classroom-based assessment (Davison, 2004; Messick, 1989; Shepard et al. as cited in Teasdale and Leung, 2000) and the socio-cultural and political factors affecting the classroom -based assessment (Davison, 2004; McNamara, 2001; Pearson, 2014).

2.4.1 Teacher assessment literacy and issues of validity and reliability

As discussed earlier in this chapter, there has been major research developments relating to educational assessment over the decades and have had some significant impact on the way that teachers think about assessment and how they implement their every day classroom assessment. The ever growing recognition of the role of assessment in the learning process has underlined the need to educate and train teachers in effective assessment practices. Webb (2002) defines assessment literacy as having “the knowledge of means for assessing what students know and can do, how to interpret the results from these assessments, and how to apply these results to improve student learning and program effectiveness.” (p.1). A wider definition is given by Fulcher (2012) who defines assessment literacy as:

The knowledge, skills and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom based tests, familiarity with test processes, and awareness of principles and concepts that guide and underpin practice, including ethics and codes of practice. The ability to place knowledge, skills, processes, principles and concepts within wider historical, social, political and philosophical frameworks in order understand why practices have arisen as they have, and to evaluate the role and impact of testing on society, institutions, and individuals. (p.125)

For Stiggins (1991), knowing how to assess also requires teachers to be well aware of their learners' learning goals and have critical thinking that will enable them to reflect on what needs to be done for effective gathering of evidence about their learners' progress. He draws attention to two main questions that teachers need to ask about their assessment tasks: (1) “What does this assessment tell students about the achievement outcomes we value? And (2) “What is likely to be the effect of this assessment on students?” (Stiggins, 1991, p. 535). These questions help teachers to use assessments methods which do not cause adverse effects but clarify the learning goals that the learners are expected to achieve. Stiggins (1991) also outlines the basic knowledge and skills that assessment literates need to demonstrate:

- Seek and use assessments that communicate clear, specific, and rich definitions of the achievement that is valued.
- Know what constitutes high-quality assessment.
- Know the importance of using an assessment method that reflects a precisely defined achievement target.
- Realize the importance of sampling performance fully.
- Are aware of extraneous factors that can interfere with assessment results.
- Know when the results are in a form that they understand and can use. (p.535)

In addition, Stiggins points out that those who are assessment literate are able to identify assessment that is not clear enough and can judge the quality of assessment to determine if it is good enough to serve its purpose.

As Popham (2009) stresses, it is important that teachers have the required knowledge and skills to be able to conduct effective assessment that promotes learning. He argues that the lack of knowledge in classroom assessment “can cripple the quality of education” (Popham, 2009, p.4) as it inhibits the learners’ progress toward the achievement of learning goals. In addition, it is believed that limited assessment literacy can raise questions of validity and reliability of teachers’ assessment. Echoing some of Stiggins suggested skills, Rogier (2014) points out that validity and reliability are some of key concepts that teachers should be conversant with in testing their learners.

Debate on validity and reliability of assessment tasks has never been a minor issue in research on educational assessment in general and classroom assessment in particular. Building on the traditional view of validity as a mere correlation between what is on a test and what students should be able to do -content/material validity- (see Ellis, 2003), Messick (1989) adds what he calls consequential validity, which is concerned with consequences of a test. Satisfying the consequential validity criteria requires test developers to take into account what consequences follow once a test is used (Chapelle, 1999). The problem in language assessment therefore is how the teacher’s classroom assessment viewed by some researchers as an “ideological” and “context-dependent process” (e.g Davison, 2004) can be well designed and carried out effectively within the classroom situation to reflect the intricacies of real life communication. Following Messick's (1989) categorization of validity, diverse views have emerged with new illustrations and/or definition of the concept of validity. One of the

suggested ways of dealing with validity issues is reiterated by Shepard et al. (as cited in Teasdale and Leung (2000) stressing that:

Performance assessments should enhance the validity of measurement by representing the full range of desired learning outcomes, by preserving the complexity of disciplinary knowledge domains and skills, by representing the contexts in which knowledge must ultimately be applied, and by adapting the modes of assessment to enable students to show what they know. The more assessments embody authentic criterion performances, the less we have to worry about drawing inferences from test results to remote constructs. (p. 164-165)

Shepard et al.'s claim emphasizes the role of clarity and variety in teachers' assessment and a need to contextualize assessment in order to collect accurate information reflecting the real knowledge level of learners. This often requires skilled and experienced teachers. Performance assessment can also be more time-consuming to design and to administer than paper and pencil based assessment (Clapham, 2000, Elliot et al. 2000). Although the day to day assessment practices can help teachers "build up a 'solid and broadly-based understanding of a pupil's attainment'" (Gipps, as cited by Teasdale and Leung, 2000, p. 165), designing effective tasks capable of providing accurate measurements necessitate well trained assessors.

However, programmes intended to train teachers on assessment are still few or even non-existent in some EFL contexts, which raises concerns over the ability of teachers to correctly assess their students (Davison, 2004). The perceived lack of skills and abilities in assessment also raises issues of reliability. Teachers' assessment is put to doubt usually on the basis that it is vulnerable to the many factors linked to the diversified environment that teachers operate in. There are questions on how classroom assessment is administered, the fairness of teachers grading and the accuracy of teachers' interpretation as well as usefulness of the results to the students' learning. In their study on formative classroom assessment practices in the UK primary schools, Gardner and Rea-Dickins (2000) found that there were still issues of validity of assessment inferences made by teachers as the reliability of their assessment was not always assured. They observed that despite teachers' evident use of formative assessment tasks, some teachers were not able to collect and interpret information on learners' language development effectively which could lead to erroneous decisions. They point out that assessment quality issues matter whether assessment is carried out for formative assessment or summative assessment purposes. For Gardner and Rea-Dickins, the multiple contexts and functions of

classroom assessment is a challenge to the possibility of establishing standards of classroom based assessment to the traditional psychometric measurement model. This warrants the need for more research to explore practices of classroom-based assessment in order to understand further the conditions of effective language learning.

Issues of appropriateness and quality of teachers' assessment can be alleviated through different ways. One of these may be to encourage teachers to collaborate to enable novice teachers to learn from experienced colleagues. Collaboration among teachers on matters of assessment can help individual teachers with finding appropriate resources and learning about the best ways to use them as teachers learn from each other (Akyel, 2000; Allal and Lopez, 2014). Collaboration can be in the form of organised discussion commonly referred to as 'collaborative groups' (Mann, 2005) where teachers regularly meet and discuss their everyday teaching experiences. It can also be in the form of 'co-teaching' where teachers prepare and deliver their lesson together to enhance learning (Liu, 2008; Stepp-Greany, 2004). Assessment literacy can also be developed through well planned training of teachers to enable them to become good assessors (Davison, 2004) and improve their students' own knowledge of assessment practices. Assessment literate teachers who use diverse and appropriate assessment methods and materials also provide the students with the opportunity to develop their awareness of the potential that assessment can offer. The students also need to have advanced assessment literacy to be able to maximize the benefits of the day-to-day assessment in their classrooms. According to Smith et al. (2013), the students need to be aware and understand the importance of the rules in their schools so that they can be able measure the quality of their work against the academic standards in place. It seems therefore that students with limited assessment literacy will not be able to participate fully in the assessment process.

2.4.2 Influence of contextual factors

In some cases, policies designed to guide classroom assessment face resistance of teachers who are slow to embrace changes and prefer to continue using their traditional methods of assessment for reasons not only related to their beliefs but also for reasons pertaining to contexts in which they practice. In his comparative study on the beliefs, attitudes and assessment practices of ESL secondary school teachers in Hong Kong and Australia, Davison

(2004) found that a divide existed between teachers regarding the practice of classroom assessment even in contexts where well established assessment criteria were present. Using teacher self report questionnaire and interview, he found that some Australian teachers were not willing to compromise their beliefs and own professional judgment for the sake of criteria-based “honesty” and “objectivity”. For some teachers, there are cases that cannot be explained by a single grade obtained from an objectively scored student coursework. This reluctance to change is caused by many factors including the lack of teacher preparation for them to understand well the fundamentals and importance of educational assessment (Cizek, 2000). Frequent and unjustified changes in school systems such as changes in examination systems can also make some teachers unwilling to shift from long practised methods acquired through experience to adopt new methods (Adamson, 2004).

In addition to teachers’ resistance to change, there are also some educational stakeholders such as parents who still undervalue results from teachers’ assessment and give higher importance to children's achievement on standardized tests. In Hong Kong where no clear assessment criteria existed at classroom level, Davison (2004) observed that different assessment methods were used in the classroom. Teachers pointed out that the existence of external examination overshadowed the value of their classroom assessment in the eyes of learners and community which resulted in indifference and little attention paid to classroom-based assessment. These results were similar to the findings in a study by Lee (2009) when she investigated the beliefs and practices of feedback of 26 teachers from Hong Kong. Her analysis of 174 documents and interview with the teachers indicated that teachers were compelled to assess and provide error based feedback to conform to the institutional norms and principles. This highlighted the degree to which the existence of contextual factors can undermine the teachers’ motivation and maintain their dependence to standardized tests instead of taking ownership and developing skills that can help them design effective classroom assessment tasks.

Important to mention is also the role of political influence on teacher classroom assessment. Teachers are sometimes compelled to do what they believe is not suitable for their respective context and act in accordance with frameworks dictated by policy makers and system managers (McNamara, 2001). The misalignment of teachers’ beliefs and their practices

is frequent in cases where policies are imposed to teachers without mutual consultations between teachers and policy makers. In her article on the change of medium of instruction in Rwanda, Pearson (2014) relays examples of teachers' statements indicating their perceived obligation to put into practice what they considered as politically and economically motivated decisions by the government. The quote from one research participant stating that "As it is a cabinet policy, we have to follow it" (Pearson, 2014, p. 44) illustrates the degree to which politically significant policies can be detrimental to the teaching and learning practices. Teachers in this environment tend to engage in any practice even when they do not agree with them and this can lead to mediocrity and ineffective assessment practices.

In actual fact, teachers often face competing demands that compel them to adjust to the reality of their teaching context. For many, the technical and practical aspect of classroom assessment present the most challenges in relation to the implementation of assessment for learning. For teachers to maximize the benefits of assessment tasks they assume a double role of being both an assessor in need of collecting sufficient evidence for reporting and accountability purposes on the one hand, and on the other hand, a facilitator with the responsibility of helping learners achieve their learning objectives. McNamara (2001) identifies three aspects of these conflicting demands: (1) the pressure to meet the assessment standards as suggested by experts, (2) administrators demands as well as (3) day to day teaching and learning obligations. He explains that teachers' responsibility to ensure validity of classroom assessment, report for accountability purposes and also guarantee that quality of teaching is maintained all affect both learning and teaching. Engaging learners in formative tasks requires teachers to find good resources fit for the purpose and this can be a daunting task for teachers, especially those at schools located in remote areas where issues of fewer resources are often reported.

Although most teachers acknowledge the importance of day to day formative assessment, the many obligations that they have to fulfill in the classroom explain the continued mismatch between what teachers claim to believe and do and what is exactly done in actual classrooms. In his study on assessment practices and instructional adjustment of 21 EFL secondary teachers in Ethiopia, Mekonnen's (2014) qualitative data suggested that teachers classroom assessment was exclusively summative oriented and assessment hardly led to instructional adjustment.

Conversely, data from self-report questionnaire suggested that assessment was used both for learning (formative) and reporting (summative) purposes. Whether these conflicting results were caused by contextual factors, by the teachers' ignorance of formative assessment or whether it was caused by the 'social desirability bias' (See Section 3.3.1), Mekonnen's findings highlight the complexity of classroom assessment practice and reemphasise the warrant for advanced research to understand further factors affecting the teachers' classroom assessment practices.

2.5 Classroom assessment and learner self-efficacy

Researchers and practitioners have long acknowledged the role that the students' beliefs play in the development of their learning motivation. Beliefs have been viewed as "strong perceptual filters" (Puchta, 1999, p. 66) that help to interpret events and predict potential performance based on such interpretations. Since the 1990s, research on the role of self beliefs in learning thrived and was expected to grow into a major educational research topic following claims in some literature that beliefs held by learners can better predict their future attainments than individual learner's skill, knowledge and prior attainments (Pajares, 1996). Some educators and practitioners considered that research on "the self" was "on the verge of dominating the field of motivation" (Graham and Weiner, 1996, p.77). Specifically, it has been argued that learners' held beliefs about their capabilities and the potential outcome of their efforts are powerful influencers of learners' actions that often lead to better achievement. For a few decades, researchers have been interested in the students' self-conception, with regards to the way that their beliefs in their capabilities affect their learning motivation.

The study of the self beliefs was particularly developed in the Social Cognitive Theory by Bandura (1986) postulating that human agency is based on reciprocal interactions between the environment conditions, personal factors and behaviour (see Figure 2.4). According to this theory, humans develop beliefs and thoughts as a result of factors such as their performance experiences and social influences based on observations and feedback from others. These beliefs influence one's behaviour as they determine one's choices of course of actions and strategies towards a certain goal, the outcome of which in turn influence the same beliefs.

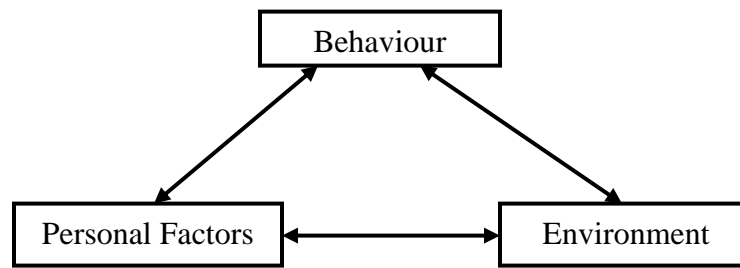


Figure 2.4: Bandura's model of reciprocal interactions (adapted from Schunk and Zimmerman, 2007)

As displayed in Figure 2.4, both personal beliefs and behaviour factors influence, and are at the same time influenced by the environment. In other words, as Pajares (1996) puts it, “because personal agency is rooted and operates within social cultural influences, individual are viewed both as products and producers of their own environment and of their social systems.” (p.544). The human beliefs, the ability to deduce the significance of performance and to decide on the course of actions to take are confounded in their capacity of self-evaluation which undergirds self related beliefs such as self-efficacy. In this study and in this section in particular, self-efficacy as a fundamental component of the social cognitive theory and widely viewed as a determinant of academic motivation is discussed to provide an overview of its tenets in the educational field. This study sought to extend research on self-efficacy in academic setting specifically in an EFL context by examining the extent to which the students’ self-efficacy is influenced by the classroom assessment. Hence, this section explores the current literature on the development of the students’ self-efficacy and the role that classroom assessment plays in enhancing learners’ self-efficacy.

2.5.1 Definition and distinction with other motivation constructs

As explained above, the self-efficacy concept stems from the social cognitive theory developed by Canadian psychologist Albert Bandura in 1980s. It is one of the most important self-beliefs theories that have attracted attention of researchers in areas of human motivation in recent few decades. Research on self-efficacy focuses on beliefs in one’s capabilities and how such beliefs influence one’s actions to attain goals. Self-efficacy theory emphasizes that what people believe is affected by what they experience and such beliefs determine how they behave and what they achieve. Bandura (1997) defines self-efficacy as the “beliefs in one’s capabilities to

organize and execute the courses of action required to produce given attainments.” (p.3). The emphasis is put on one’s beliefs rather than one’s demonstrated competence to predict the future attainment. According to self-efficacy theory, people who believe that they have the capacity to achieve are able to see difficulties as opportunities, learn from experiences and use them to predict the outcomes of their actions.

Self-efficacy is also distinguished from other closely related concepts. These include mainly terms such as ‘self-concept’, ‘self-esteem’, ‘perceived control’ and ‘confidence’. While self-efficacy is about the perceived abilities to perform a specific task, *confidence* is defined as a “nondescript term that refers to strength of belief but does not necessarily specify what the certainty is about” (Bandura, 1997, p. 382). Put in an academic context, a student may be confident that she or he will fail on a test in general. Another related construct that is also generic is *self-concept*. This is generic in that it does not relate to a specific task in a given domain. It is “one’s collective self-perceptions that are formed through experiences with, and interpretations of the environment, and which are heavily influenced by reinforcements and evaluations by significant others” (Schunk and Pajares, 2002, p. 16). Schunk and Pajares argue that self-concept is centred on feelings of self-worth based on one’s held beliefs of competence. Having low self-concept may reduce one’s interest or commitment to a task due to the absence of self-worth attributes associated with the task. In a language class, the development of students’ high self-concept may be enhanced or undermined by what happens in the classroom particularly when students are not well supported in their effort to communicate in the target language (Arnold, 2007). High self-concept is better developed in classroom environment that is free from practices such as “labelling, criticism, sarcasm, put downs, comparisons, and evaluating the person rather than the behaviour” (Arnold, 2007, p.18)

Self-esteem is another motivation construct that is different but also related to self-efficacy. Similarly to self-concept, self-esteem refers to the personal sense of worth. It is “the evaluation which the individual makes and customarily maintains with regard to himself [...], a personal judgment of worthiness that is expressed in the attitudes the individual holds towards himself” (Coopersmith, 1967, 4-5). Unlike self-efficacy, self-esteem is not domain specific and may develop throughout the entire life span of a person (Zimmerman, 1995). On the other hand, *perceived control* is defined as “generalized expectancies that the self can

produce desired and prevent undesired outcomes (similar to expectancies of success)” (Skinner and Greene, 2008, p. 6). Perceived control is about causal attributions of performance outcomes and rests on the belief that people develop awareness of what is required for them to perform a task and their belief that the required skills are available and can be used to reach the desired outcome (D’Ailly, 2003). In education context, students with high perceived control are likely to believe that their success will come as a result of their high efforts and that they are able to exert such efforts for high achievement.

2.5.2 Sources of self-efficacy

Bandura (1997) names four main sources of self-efficacy beliefs as (1) enactive mastery experiences, (2) vicarious experiences, (3) verbal persuasion, and (4) the physiological and affective state of an individual. While these are commonly cited as contributing factors to the development of self-efficacy, the enactive mastery experience has been often considered as the major source of self-efficacy in academic settings (Pajares, 1996). The students’ experience of success and its attribution has been viewed as key influencing factors to their self-efficacy and subsequent academic performance. However, it is also important to note that some educational psychologists have suggested that academic success does not necessarily depend on positive self beliefs but on the student’s goals and perception of failure and success (Dweck, 1999). According to Dweck’s implicit theories, students who believe that intelligence is malleable (Incremental theorists) pursue learning goals. They perceive achievement as product of effort, learn from mistakes and they are not weakened by failures.

Conversely, students who think that their intelligence is fixed (Entity theorists) are performance-oriented and failures are debilitating to them. Dweck’s implicit theories underline a multidimensional aspect of self-beliefs and highlight other possible cognitive factors that lead to learning motivation and achievement. Nonetheless, the link between students’ self-efficacy and their academic performance has been widely considered as a fact with the students’ experience of success as an influencing factor to the students’ motivation to work with effort and perseverance in their courses (Brookhart, 1998; Zimmerman, 1995).

2.5.2.1 Enactive mastery experiences

Bandura (1997) explains that experience of success i.e. “enactive mastery experiences’, constitutes one strong foundation of a person’s self-efficacy beliefs when successful accomplishment of a task is considered by the performer as a strong indicator that he or she is capable of attaining his or her objectives in that specific task. However, while Bandura and other self-efficacy theorists point to success as the most powerful source of self-efficacy, they also indicate that perceived failure to complete tasks successfully can lead to the decrease of positive self beliefs especially when much attention is put on failure and less on success. In academic setting, the relationship between learners’ self-efficacy and academic success has been debated with some views maintaining that experience of success leads to strong self-efficacy whereas other views consider success as a result of self-efficacy (Pajares, 2001). The latter emphasize that self-efficacious students have high academic attainment because they are motivated to learn, provide effort and do not give up on challenging tasks.

Bandura’s (1978) concept of ‘reciprocal determinism’ highlights that there exists a reciprocal influence between self beliefs and achievement. He argues that self beliefs influence and are influenced by human behaviour and environmental factors, maintaining that personal experiences is the most dominant source of self-efficacy once they are considered genuine, real evidence of one’s competence required to execute a task successfully. The perception of success as a merit is important for the growth of positive self beliefs. Talking of the importance of authenticity of mastery experiences, Erikson Erik (1994) states that even children feel fully positive about themselves and develop strong self-esteem only when they experience “wholehearted and consistent recognition of real accomplishment, that is, achievement that has meaning in their culture.” (p. 95). Authentic evidence of success becomes a motivational factor to engage in future tasks, which in turn removes the fear of difficulty, anxiety and other factors that often prevent people from engaging in new activities.

In addition to the significance of authentic evidence as an essential requirement for the increase of self-efficacy, Bandura (1997) also indicates that the increase or decrease of self-efficacy may as well depend on factors such the student’s preconceived capability, perceived task difficulty and effort made, the amount of external aid that the student gets or the

circumstances under which success or failure occurs. Students bring with them a preset image of themselves when approaching a task and that can play a leading role in attributing success or failure to the amount of effort or level of capability.

2.5.2.2 Vicarious experiences

Vicarious experiences strengthen self-efficacy beliefs by providing real evidence through models which provide evidence that a successful completion of a task is possible. This evidence is obtained by observing individuals modeling tasks and completing them successfully (Bandura, 1995). If the observed people have a lot in common with the observer, the success will be exclusively attributed to the competence of the performer and not to other factors. This strengthens the observer's beliefs that they can achieve the same success considering that the person who achieved is not different from them. This is described as 'observational learning through modeling' by Schunk (2003) who argues that observing others perform can be either motivating or demotivating depending on the rewards or punishment associated with the performance. When the modeled performance leads to rewards, the learner's self-efficacy will be raised and will encourage future performance. However, vicarious experiences alone cannot be relied upon for the development of a person's self-efficacy. It has been argued that observing other people performing well can be debilitating to some observers who may consider other performers' success as a debilitating factor (Dweck, 2007). According to Dweck, such personal beliefs are found in 'Entity theorists' who pursue performance goals in their tasks and attach much importance to performance rather than learning. Dweck stresses that performance models are more beneficial for 'Incremental theorists' students who are learning-oriented.

Using data from a longitudinal study on self and personality development of 508 college students in California, Robins and Pals' (2002) validated Dweck's self-theories in their study on academic implications of self-theories for students' goal orientation, attributions, affect, and self-esteem change. They found that students with Entity theory beliefs tended to adopt performance goals whereas those with Incremental theory beliefs tended to opt for learning goals. The study indicated that students with mastery goal orientation attributed their failure to the insufficiency of effort and reported learning from others while those with performance goal orientation perceived their failure amidst success of others as a proof that they lacked abilities.

The feeling of unease in time of failure and failure to learn from others' success is often found in learning environments where product (e.g. grades) is more valued than process thus giving way to competition and learners pursuing performance goals rather than mastery goals (see Alkharusi, 2008). In these contexts, vicarious experiences may lead to the decrease of positive self beliefs among low performing students.

2.5.2.3 Verbal persuasion and individual physical and affective state

Self-efficacy beliefs can also be strengthened through encouragement by other people through verbal persuasion. Verbal persuasion is the attempt to get a person to do a task by verbally convincing him or her that he or she has the competence to do it successfully. Although verbal encouragement may be most effective in increasing self-efficacy in some circumstances, Bandura (1997) indicates that it is the least influential of the four other sources of self-efficacy. In actual fact, it may be hard to convince a person who thinks he or she knows the limit of his or her own capabilities, that he or she is capable of accomplishing a task. Verbal persuasion usually requires supplementary reinforcement inputs and this is where self-efficacy meets classroom assessment on assumption that what students get as feedback on their performance matters for their self-efficacy development. In formative classroom assessment, students get teachers' and peer's feedback that may be offered in form of encouragement which helps learners to develop confidence (Schunk, 1982).

The encouragement for future success should be based on the accomplished success especially with experienced learners (Ritchie, 2016). On the other hand, Bandura (1997) indicates that the way that individuals interpret their physiological and affective states can have significant influence on their self-efficacy. Personal feelings of physical strength and good health may serve as a good sign that one has the capacity to perform a task while the feeling of physical weakness and emotional distress can be interpreted as a sign of physical and affective incapability (Bandura, 1997). In the academic setting, the students' physiological and affective states may include emotional states such as fear and anxiety that make students feel vulnerable for poor performance. However, the positive mood is also considered as a booster of self-efficacy as students will feel in good position to perform well.

2.5.3 Effects of self-efficacy on learning motivation and achievement

Studies in language education and in educational research in general have provided evidence of a positive relationship between self-efficacy, learning motivation and academic achievement respectively (Hsieh, 2008; Mills and Pajares, 2007; Multon, Brown, and Lent, 1991; Zimmerman, Bandura and Martinez-Pons, 1992). From a social cognitive theory perspective, Learners with a high level of self-efficacy are expected to be more motivated than students with a low level of self-efficacy. Perceived self-efficacy determines the level of persistence, choice of activities and level of effort provided by the learner (see Schunk and Pajares, 2002). Research has also indicated that learners' future academic achievement is better predicted by the level of learners' self-efficacy than affective factors or prior performance. Using a 21 self-efficacy item measure constructed by the researcher, Chen's (2007) survey investigated the predictive power of learners' self-efficacy, anxiety and perceived value of English language and culture on 277 Taiwanese college students and the Bivariate and hierarchical multiple regression analysis indicated that self-efficacy was the best predictor of students' performance in listening.

Chen's findings agree with results from a meta-analysis by Multon et al. (1991) on the relationship between students' self-efficacy and their academic performance and persistence. The analysis pointed to more evidence that a significant relationship existed between learners' self-beliefs and their achievement not only across levels of studies but also across subjects. This was also echoed in Zimmerman et al. (1992) survey on the role of self-efficacy beliefs and goal setting on academic attainment. Their study involving 103 ninth and tenth American graders, found that the learners' self-efficacy had a significant influence on their achievement mediated by learners' goals. Using path analysis methods, they also noted that other self-motivational variables such as learner self-regulation were linked to learner performance and were predictive of self-efficacy for academic achievement. As displayed in Figure 2.5, Schunk (1995) presents a model highlighting the process through which self-efficacy mediates the influence of personal qualities, performance experience and social influences and motivation. The two stages of self-efficacy shown in the model demonstrate a continuous process that starts with learners' beliefs about and perceptions of their abilities influenced by factors such as support from important people in their environment.

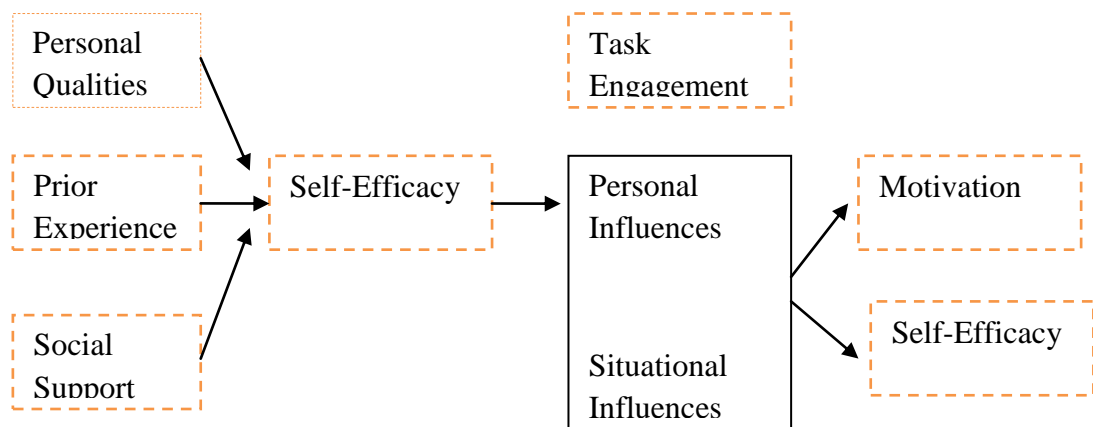


Figure 2.5: the role of self-efficacy in academic achievement (adapted from Schunk, 1995)

The support may come in the form of feedback from teachers on their performance or encouragements and guidance from their parents. Such encouragements increase the students’ beliefs that they can perform tasks and motivate them to engage in tasks. Success in tasks leads to the feeling of satisfaction and confidence which can in turn strengthen the learner’s self-efficacy.

In the context of foreign language learning, there has been relatively little progress in research on the influence of self-efficacy on learning motivation and achievement of EFL learners. However, some research studies and theoretical works have pointed to the influence of self-efficacy and other self-beliefs on foreign language learners’ achievement (Naseri and Zeferanieh, 2012; Mills, Pajares and Herron, 2007; Pui, 2010; Zimmerman and Bandura, 1994). In their survey of 303 US college students learning French as a foreign language, Mills et al. (2007) observed that a relationship existed between learners’ self-efficacy for obtaining good grade and self-efficacy for self-regulated learning which was also found to be a strong predictor of French learners’ achievement. These findings confirm Bandura’s (1997) view that self-efficacy of learners is associated with their self-regulated learning strategies where self-efficacious learners adopt effective organisational and functional strategies to achieve their academic goals. Mills and Pajares’ study results are also consistent with other study findings that have demonstrated correlation between learners’ self-efficacy and perceived ability to strategize and their perceived ability for high achievement. In their study involving 80 EFL

Iranian learners, Naseri and Zeferanieh (2012) used the Reading Self-efficacy Questionnaire (0.81 Cronbach's Alpha) adapted from Ghonsooly and Elahi (2011) to investigate the relationship between the learners' reading self-efficacy and reading comprehension. The results of their Spearman correlation showed a statistically significant and positive correlation between self-efficacy beliefs in reading and high levels of reading comprehension.

2.5.4 Influence of assessment on learner self-efficacy

In view of the discussion above on the sources of self-efficacy, there seems to be evidence indicating that what happens in the classroom in terms of assessment can be critical in shaping students' self-efficacy beliefs. If the classroom is considered as a micro social context one may expect the learners to be influenced by attitudes and behaviours of both the teachers and the peers as Taylor (2013) explains:

Filling a large proportion of the adolescents' time, the classroom is a micro social setting that leaves its socio-ideological mark on students' identity through the mediation of teacher beliefs and practices [...]. In addition, given that students tend to perceive the teacher's responses as assessment of themselves as persons rather than of their performance, the feedback given in class is also crucial: not only should it be informative rather than controlling, but it should emphasise effort rather than ability or intelligence. (p. 15)

In addition to the teacher's feedback, Taylor also points out that peer judgments expressed through socialisation at school are very crucial to the development of learners' self beliefs and academic orientation. Empirical studies show that through students' self-assessment and timely and effective feedback from peers and teachers, learners get opportunities to appreciate their efforts and their accomplishments (e.g. Belegizadeh and Masoun, 2013; Coronado-Aliegro, 2006). The feeling of satisfactory achievement affects the student's perceptions of competence, increases their self-confidence and their self-esteem (Oroujlou and Vahedi, 2011), which in turn influences subsequent performance. Kuciel (2013) explains further the link between assessment and self-efficacy stating that "...as self-efficacy is postulated to be chiefly the product of mastery experience, then students' perceptions or experiences of past performances in the foreign language can be crucial for establishing high self-efficacy." (p. 37). Conversely, it can be expected that unsatisfactory performance may be detrimental to learners' self-efficacy as they may develop beliefs that they cannot perform to the desired level.

In his quantitative study on the effects of continuous self-assessment on the self-efficacy of 104 undergraduate American learners of Spanish as a foreign language and using the task and skills-specific item self-efficacy scale adapted from Mills (2004), Coronado-Aliegro (2006) analysis of covariance (ANCOVA) found that learners who were continuously engaged in self-assessment throughout the semester felt more self-efficacious than learners who had had less experience in self-assessment. Their involvement in assessment activities helped them develop self-assessment skills that enabled them to rethink and understand further the learning goals and to adopt effective alternatives enabling them to reach those goals. Although there is limited literature on the influence of classroom assessment on the learners' self-efficacy in EFL context, the link between assessment practices and self-efficacy development has been identified in studies from other educational fields particularly in sciences education research (e.g. Alkharusi et al., 2014; Dorman et al., 2006). In their qualitative study of on 'the influences on student cognitions about evaluation', Ross, Rolheiser and Hogaboam-Gray (2002) investigated the beliefs of 71 students from Multilanguage grade 2, 4 and 6 classes in Canada on classroom evaluation and how it affected their self beliefs. Their theme analysis of interview data found that the students' beliefs and knowledge about assessment contributed to the development of their self-efficacy. Such beliefs were found to be a result of peer and particularly parental influences. The findings indicated that students whose parents were involved in the identification of achievement standards and in the raising of their aspirations demonstrated strong self-efficacy and for them effort was seen as path to success.

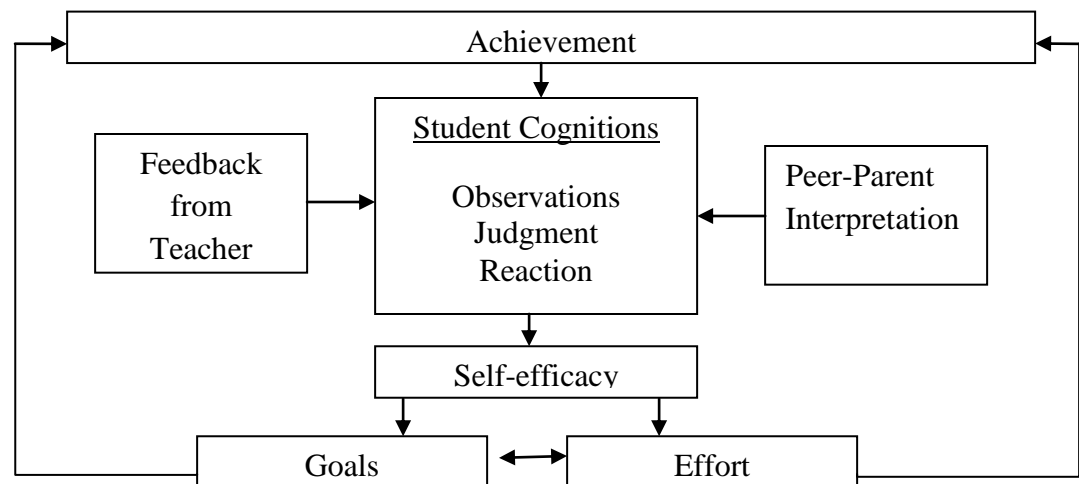


Figure 2.6: Effect of classroom assessment on the student's achievement by mediation of student's self-efficacy (adapted from Ross et al., 2002)

As displayed in Figure 2.6, Ross et al. (2002) describe the teacher feedback and parental and peer influences as key contributors to the development of self-efficacy. Ross et al. indicate that learners self-efficacy is promoted through assessment when feedback obtained by learners help them see their weaknesses and strengths which shapes their beliefs about themselves. In this cycle, 'feedback' is shown as being at the heart of the relationship between teachers' assessment and students' self-efficacy where assessment methods that actively involve learners and help them reflect on their own learning are expected to strengthen learners' self-efficacy (Mills, 2013). Like teachers' feedback, peer progress feedback that provides evidence of success also constitutes an environmental variable that influences learners' self-efficacy (Schunk and Rice, 1991; Schunk and Zimmerman, 1997). Effective feedback is critical for raising learners' awareness of their own learning progress and the awareness of own weaknesses and strengths can enhance learners' engagement in learning tasks and can become a source of commitment to achieve improvement. However, it is also important not to focus feedback on learners' weaknesses as this may prevent learners from exerting their effort for better improvement. At the sametime, the learners' self-efficacy needs to be well 'calibrated', i.e. appropriately judged to accurately reflect the real learners' abilities (Phakiti, 2006).

According to Hattie and Timperley (2007), effective feedback addresses three questions: 1) *where am I going?* (feed up dimension) that seeks to get information on the learning goals that have to be attained, 2) *How am I going?* (Feedback dimension) which informs teachers and/or learners about their progress towards attaining their goals and 3) *Where to next?* (feed forward dimension) that helps learners identify what is needed for achieving more. Emphasising more on the feedback dimension, Hattie and Timperley indicate that feedback can have four levels: 1) feedback directed to the task (FT) i.e. whether it is correct or wrong, 2) feedback about the process used to complete a task(FP), 3) about the learner self-regulation (FR) such as the learners' confidence to take up further tasks and 4) feedback directed the learner's "self" (FS) where teachers may utter praise words such as 'you are a great student'. The FR and FP are considered the most influential for achievement as they reemphasise what is needed for improvement and can raise the learners' self-efficacy and encourage subsequent deep processing and mastery. However, Hattie and Timperley also indicate that integrated levels can be better and argue that feedback that praises students efforts and acknowledges their

knowledge and capacity to successfully complete a task on their own "can have major influences on self-efficacy, self-regulatory proficiencies, and self-beliefs about students as learners" (Hattie and Timperley, 2007 p. 90).

In general, positive feedback often acts as a promoter of students' positive beliefs about themselves and this can increase their motivation to learn. As Pajares (2003) explains, the influence of self-efficacy on motivation is based on the feelings and beliefs that students develop as a result of their interpretation of performance as successful, which boost their confidence. Students who judge their performance as successful are likely to develop high internal desire to do further tasks (intrinsic motivation), pay more attention, put more effort in their work and demonstrate high level of perseverance and resilience when they encounter challenges. It can be therefore assumed that engaging students in tasks where they are likely to experience success, whether reinforced by their own judgement or by feedback from peers and teachers, will raise the students desire to do more and be able to face up challenging tasks. The tasks also need to provide opportunity for real-life like performance to reflect real abilities of the students and not solely driven by external rewards that may only produce short term interest. In his study on the effects of teachers' assessment practices on secondary school students' achievement goals in science classroom in Oman, Alkharusi (2008) noted that different forms of assessment and the ways they were used in the classroom had a significant impact on students' learning goals, their motivation and the eventual achievement outcomes. Using the hierarchical linear modeling techniques to analyse data from 1,636 ninth grade students and 83 science teachers from public schools in Oman, Alkharusi's study results indicated that assessment using 'traditional' methods promoted the learners' extrinsic motivation even for students with high self-efficacy levels. Students were found to be vulnerable for having "motivation driven by external rewards (extrinsic) instead of being driven by internal desire to achieve mastery (intrinsic motivation). This in turn encouraged students to study for other purposes such as gaining good grades other than mastering and understanding the content.

On the other hand, the use of less formal and less controlled methods of assessment - alternative assessment such as writing research papers or engaging in oral performances- has been found to increase the students' motivation when students are successful in their tasks. By

experiencing success, the students are likely to feel more confident. This motivation enhances persistence of learners in case of difficulty during their tasks and encourages them to take up challenges in their study, which is characteristic of learners who feel self-efficacious. While Alkharusi's (2008) findings agree with other previous study results, that students' self-efficacy is higher in classroom assessment contexts where less controlled and less formal methods are frequently used, it is important to consider that his research relied on data collected solely by means of a survey that may not suffice for a good grasp of what really happened in the classroom. Classroom is a complex social setting that research on classroom practices requires the use of effective and carefully selected methods to explore fully the intricacies of the classroom assessment practice.

2.6 Summary of the Chapter

Chapter two has explored the current highlights in the literature on the classroom language assessment and the learner's self-efficacy in EFL context. The chapter first presents a brief historical overview of the English language testing and identifies the significant role that the testing practices particularly the use of standardized and high-stakes tests have played in shaping the teachers' classroom-based assessment. The disenchantment with the use of such tests is highlighted and presented as a basis for the renewed call for the effective classroom assessment practices that allow learners to participate in the assessment process in order to develop better understanding of the learning potential that assessment can provide. Second, the chapter equally explores the literature on the development of learners' self-efficacy and examines the significant role that it plays in improving learning achievement. The literature argues that self-efficacious learners are likely to strategize when faced with challenges. They are motivated to do more and eventually achieve high as a result of their confidence and persistence. While the current reviewed literature indicates that the classroom assessment and the learner self-efficacy both play an important role in improving learning, it is stressed that little effort has been made to examine the relationship between these two variables. Some literature is discussed to highlight the importance of investigating such a relationship and this underpins the rationale of the current study.

Chapter 3: Methodology

3.0 Introduction

This chapter is an overview of the underlying research principles that have guided the design and the implementation of the current study. It presents a brief review of the theoretical assumptions and the epistemological basis that led to the formulation of the research questions, the sampling techniques and procedures, the data collection instruments and process as well as the methods used for data analysis. A mixed method approach was used to investigate what teachers did to assess their students in the classroom, to determine the level of the students' self-efficacy across the four English language skills and to examine the extent to which the students' self-efficacy was related to the teachers' assessment practices. This chapter also summarises the background data of the respondents in terms of the demographic information and other data that are essential for full description of the respondents. The background presented in this chapter consists mainly of the gender and age of students and teachers, the status and learning mode of every school that participated in the study. The background data also provides more details on the teachers' in-service training and professional qualification. The chapter also discusses other key issues that were taken into consideration before and/or during data collection. These include a discussion of ethical principles that guided the process of data collection.

3.1 A mixed methods approach

The researcher in the current study sought to understand the practice of classroom assessment and its relationship with the students' self-efficacy in the Rwandan lower secondary school by collecting and analysing data using a variety of methods and instruments. Data relating to each research question was gathered from different respondents in different ways (See appendix 12). In other words, the researcher adopted a pragmatic approach in the search of answers to the research questions. The ontological and epistemological dualism of subjectivity and objectivity was viewed as complementary rather than incompatible paradigms. Therefore, instead of entirely depending on one single paradigm, the mixed methods approach was found to be

appropriate for the current study. The researcher took the stance that truth and reality can take multiple forms and can be both objective and socially constructed (Denscombe, 2008; Johnson and Onwuegbuzie, 2004). As a result, the current study was intended to produce more literature on EFL classroom assessment and learners' self-efficacy derived from both qualitative and quantitative based research. Many of the previous studies in the area of classroom assessment and self-efficacy have often used either quantitative methods (e.g. Belegizadeh and Masoun, 2013; Deluka et al. 2016; Dunn, Strafford and Marston, 2003; Inbar-Lourie and Donitsa, 2009; Jenks, 2004; Tsagari and Pavlou, 2009; Yang, 2008) or qualitative methods (e.g Cheng et al. 2004; Chumun, 2002). A few other studies have also used mixed methods (e.g. Mekonnen, 2014; Mussawy, 2009; Zimbicki, 2007) that have led to important literature.

The researcher considers that there is a need for more research studies using the mixed methods. The current study thus uses both qualitative methods and quantitative methods as the most effective way of achieving rich data. This paradigmatic flexibility facilitates the collection of both numerical and qualitative data which help to address the 'what' and 'how' types of research questions (Cohen, Manion and Morrison, 2007). Both qualitative and quantitative mixed methods were thus used in this study to collect data guided by two main research questions that included both 'what' and 'how' questions (see Figure 3.1). It is a *concurrent triangulation* design where both quantitative and qualitative data were collected at the same time and information integrated in the interpretation of the overall results (Creswell, 2003). The qualitative methods were informed by the quantitative methods where the respondents' participation in the process of quantitative data gathering yielded eligibility to participate in the qualitative part.

It is also a correlational design that investigated the degree of relationship between variables (Creswell, 2005; Martella et al. 2013). In other words, it was conducted to identify the extent to which teachers' classroom assessment practices, particularly the use of forms of assessment and feedback, were related to the students' self-efficacy for the reading, listening, writing and speaking of English. The main stages of the study included designing data collection instruments, piloting and administering the instruments, organizing and analyzing both quantitative and qualitative data, presenting and discussing the results as well as writing conclusions. The data collection methods consisted of the classroom observation, the teachers'

interview and students' focus groups. In addition, both teachers and learners of English at the lower level of secondary schools participated in a questionnaire survey. The process of data collection was done after the early stages were completed i.e. the review of literature on classroom assessment and the social cognitive theory of which self-efficacy is the main component. This was discussed to provide the theoretical foundation upon which this study was built and to link this study to the relevant and most recent study findings (Hammond and Wellington, 2013; JHA, 2014).

3.2 Research questions

This study aimed to explore the teachers' classroom assessment practices and how these practices were related to the students' self-efficacy in the four English language skills at the lower level of secondary schools in Rwanda. To attain this objective, the investigation was guided by two major research questions. The first research question sought to explore the classroom assessment practices of EFL teachers at the lower level of Rwandan secondary school and the students' perceptions of such practices. The focus was put on the forms of assessment and methods of providing feedback that the teachers used in their classroom. Research question one also explored other aspects of the teachers' classroom assessment such as the purpose and time of assessment as well as sources and content focus of assessment tasks. The overall students' perceptions of the teachers' assessment practices were also examined under research question one. The second research question investigated the students' self-efficacy across the four English language skills and the extent to which such self-efficacy was related to some of the teachers' assessment practices described in research question one.

In order to collect the necessary and manageable data that is relevant to the study, the two research questions and their corresponding sub-questions were constructed in strict consideration of the study objectives (Andrews, 2003; White, 2009). One of the research objectives was to explore and describe the assessment practices specifically what teachers did in the classroom in relation to the ways by which teachers gathered information about their students' learning progress and how they provided feedback. In their everyday classroom assessment practices, teachers use different ways of assessment to gather information about the teaching process and/or about the students' learning progress.

Research question (RQ)	Methods of data collection
<p>RQ1: What are the classroom assessment practices of teachers of English and how do the students perceive these assessment practices?</p> <p><i>RQ 1.1 What forms of assessment do the teachers use in their classroom?</i></p> <p><i>RQ 1.2 For what purpose and at what time do the teachers use assessment forms in the classroom</i></p> <p><i>RQ 1.3 What is the content focus and source of assessment tasks used in the classroom?</i></p> <p><i>RQ 1.4 What methods of providing feedback do the teachers use?</i></p> <p><i>RQ1.5 What are the students' perceptions of the teachers' classroom assessment practices?</i></p>	<p>Questionnaire, classroom observation, Interview</p>
<p>RQ2: To what extent is the students' self-efficacy for the four English language skills related to the teachers' assessment practices?</p> <p><i>RQ 2.1 What is the level of the students' self-efficacy for the four English language skills?</i></p> <p><i>RQ 2.2 How is the students' self-efficacy for the four English language skills related to the students' perceptions of the teachers' assessment practices?</i></p> <p><i>RQ 2.3 How is the students' self-efficacy for the four English language skills related to the teachers' assessment practices</i></p>	<p>Questionnaire, classroom observation, interview</p>

Figure 3.1: Research questions and methods of data collection

Teachers may use 'general ways' such as learner observation or learner questioning, or the 'specific ways' such as oral presentations and essay writing (Butler and McMunn, 2006). In this study, the teachers' ways of assessment were described and classified in two forms: *paper and pencil* forms for assessment that uses true/false items, fill-in-blanks, multiple choices or short answers, and *performance assessment* forms for assessment such as oral presentations, interviews, and debates (see Section 2.2.4). Answers to sub-research question 1.1 yielded data

on teachers' most preferred forms of assessment between *paper and pencil*, and *performance based* as well as information on the methods that teachers used to provide instructions regarding assessment (use of rubrics). As shown in Figure 3.1, the sub-research questions 1.2 and 1.3 were formulated to guide the collection of data on other teachers' assessment routines while sub-research question 1.4 was centred on data about the teachers' methods of providing feedback to their students.

As feedback is an important component of assessment, information was collected regarding the different methods of providing feedback that teachers used in their classroom. Sub-research question 1.5 was focused on data about the students' perceptions of the classroom assessment for further understanding of the mode of influence between the students' self-efficacy and the teachers' assessment practices. Data collected for research question two was used to examine the level of the students' self-efficacy for English speaking, reading, listening and writing (sub-research question 2.1). It was also used for deep analysis of the relationship (1) between the students' self-efficacy and the students' perceptions of the teachers' assessment practices (sub-research question 2.2), and (2) between the students' self-efficacy and the teachers' assessment practices (sub-research question 2.3).

3.3 Methods of data collection

Four methods were used to collect data for this study: The questionnaire survey for teachers and students, the interviews for teachers, the students' focus groups and the classroom observation (See appendix 12 for methods and data collected for each research question). The teachers' interviews, the students' focus groups and classroom observation were used to supplement the information obtained by use of questionnaire survey.

3.3.1 The questionnaire survey

The questionnaire is one of the common data collection instruments especially in social sciences research where it is often used as the source of primary data. It is usually designed to gather structured and numerical information about people's life conditions, their attitudes and their held beliefs about a given topic (Murray, 1998; Cohen et al., 2007). Although a questionnaire can take many different forms to fit different data collection purposes, it is

usually classified in three main categories depending on the classification of question items on a continuum from highly structured questionnaire, semi-structured and highly unstructured questionnaire. Question items in a highly structured questionnaire are constructed before its administration and are often expected to elicit numerically analyzable finite data (Pathak, 2008). These question items are also often classified as either *closed questions* where respondents are required to choose their answers to the asked question from a provided list of answers or *open questions* where the informants' responses to questions are presupposed but not explicitly stated in the questionnaire (Cohen et al., 2007). On the other hand, the semi-structured questionnaire indicates the focus of the questions but does not presuppose the informants' responses. Informants are free to write what they want in their own words but limiting their information on the agenda set by the researcher. This differs from the highly unstructured questionnaire where respondents are invited to write what they want and are not necessarily required to limit themselves to the researcher's agenda.

The use of a questionnaire in this study was beneficial in that it allowed for the collection of information from a large number of respondents in a relatively short time. According to Kothari (2004), a questionnaire also helps reduce the interviewer bias and allows time to the respondents to think about their responses. In most cases, the respondents in the current study were given the questionnaire during their free time or during the weekend when they had no other curricular activities. This allowed them sufficient time and freedom to think about the questions. Another advantage of using the questionnaire is that it does not reveal any identifiable information concerning informants and this can increase honesty in the respondents' answers (Cohen et al., 2007). Although participants in this research were not asked to disclose sensitive personal information, it was expected that some respondents might not want to reveal their identity while commenting on their own beliefs or on their teachers' classroom assessment routines. Therefore, the questionnaire had a particular advantage of keeping the respondents anonymous (Solomon, Marshal and Stuart, 2006). The questionnaire is also chosen because, as Cohen et al., (2007) emphasize, it enables to elicit information that is finite and easy for analysis.

In spite of the advantages outlined above, the use of a questionnaire in survey based research is not without shortcomings. One of these is the possibility of low return rate of the

questionnaire which can become a major challenge if data obtained is not enough for the purpose. In addition, it has been long argued that data collected by use of self-reported questionnaire can be subject to the “social desirability bias”. This is defined as “...a bias resulting from participants giving responses that make them look good rather than giving honest responses” (Mitchell and Jolley, 2010, p. 629). The respondents may prefer to provide information that is likely to project their own positive image instead of relaying the real information. This bias may also occur when participants in the survey report information that is widely and socially accepted or when they believe that the researcher would be more receptive to the information (Fisher, 1993; Krumpal, 2013; Marsden and Wright, 2010). Although it may be practically impossible to remove the possibility of this bias, there are a number of strategies that are recommended to help reduce the respondents’ propensity to giving false information. These include mainly emphasizing the anonymity aspect of the collected data as a measure to protect the privacy of the informants. Except in the case of ‘unconstructive measurements’ where respondents are purposely recorded without their knowledge (Mitchell and Jolley, 2010), the respondents need to be provided with clear information about the purpose of the research and be assured that their information will be protected (Cheng and Dornyei, 2007). It is also possible to reduce the social desirability limitation by providing the respondents with many options for their answers whenever possible to allow some degree of freedom and feeling of control.

3.3.1.1 Structure of the student questionnaire

In the current study, a semi-structured questionnaire was used to collect data from all informants who were involved in the investigation i.e. teachers and students of F2 and F3 of lower secondary schools in the selected districts of Rwanda (See appendix 1A). Also, while 3% of the students’ questionnaire items were open ended-questions that were designed to give the respondents the opportunity to add more clarifications in their own words, up to 97% of the question items in the questionnaire were formulated to yield quantifiable data easily manageable and analyzable manually and/or by using data management and analysis software programmes. On the one hand, the respondents were presented with a number of statements from which they had to choose those that best described their point of view. On the other hand, however, for other questions, the respondents were required to write their extended answers

which were recorded and later analyzed as provided. All question items relating to each theme were grouped together and presented in the same format to facilitate the respondents' easy understanding of the items (Gravetter and Forzano, 2015).

Table 3.1: Structure of the questionnaire for students

Part	Theme and content	Number of Items	Scale
I	1. Demographic information	5	-
	<i>-Age, gender, home residence, average grade, parents' level of education</i>		
II	2. Teacher classroom assessment practices	8	4-point Likert-scale
	<i>-Congruence, Authenticity, Student consultation, Transparency, Diversity, Frequency, Feedback and teacher fairness in marking</i>		
	<i>-Forms of assessment</i>	6	
III	3. Perceived self-efficacy for speaking English	5	11-Point Scale
	<i>-Speaking outcomes for F2 and F3 (for students in year 2 and 3 LL)</i>		
	4. Perceived self-efficacy for writing	5	
	<i>-Writing outcomes for F2 and F3</i>		
	5. Perceived self-efficacy for reading	5	
	<i>Reading outcomes for F2 and F3</i>		
	6. Perceived self-efficacy for listening	5	
<i>Listening outcomes for F2 and F3</i>			
Total items		39	

Question items were written in a logical order to allow the smooth flow of ideas and good transition from one theme to the other (Sarankatos, 2013). They were well sequenced from simple to complex and from closed ended to open ended (Wellington and Szczerbinski, 2007). Question items were also formulated in such a way that the generated answers would be easily recordable and analyzable with the planned data analysis method. It is in this regard that question items in the student questionnaire were organized in five content themes arranged in three parts (See Table 3.1). Theme one was about the student's demographic information in part one, theme two was the student's perceptions of classroom assessment practices in part

two, and part three was comprised of theme three on the student's perceived self-efficacy for speaking, theme four about the student's perceived self-efficacy for writing, theme five on the student's self-efficacy for reading and theme six on the student's self-efficacy for listening. The question items were written in clear language and were distinguishable from each other to avoid confusion or skipping of questions.

As described in Table 3.1, Part one contains 5 items on demographic information. This concerns the basic facts on the identity of the respondent (Goodwin, 2009). These questions were designed to collect data on the student age and gender, their home residence status, the grades average in English for one previous term and the parents' level of education. According to Colton and Covert (2007), information on demographic variables can be useful in determining the context of the respondents' answers and in judging the respondents' representativeness of the population. In the case of this study, the demographic data provided information about the degree to which the participating respondents reflected the reality of the population for example in terms of mixed gender of the secondary lower level which is nearly 50% male and 50% female or in their age range between 13 to 18 (MINEDUC, 2015b).

The 15 items in part two were formulated to elicit information on students' perception of the teachers' assessment practices in relation to congruence of assessment tasks with learning content, transparency and authenticity of the tasks, frequency of assessment, diversity of assessment methods, student consultation and involvement in the classroom assessment and teacher feedback and teacher fairness in marking. To express their perceptions, the respondents were asked to choose the extent to which they agreed or disagreed with statements on a 4-point Likert scale of *Strongly Agree=4, Agree=3, Disagree=2, and Strongly Disagree=1*. In addition, students had to provide their answer to one open question in this section. Question item ten in part two (See appendix 1A) sought to gather information from the students' perspective on how frequently their teachers used different forms of assessment.

Part three contained 20 question items organised in four themes relating to students' self-efficacy. These question items were designed to investigate the students' self-efficacy for performing different tasks in speaking and writing of English as well as in reading and

listening. The respondents measured their answers against an 11-point scale from 0 to 100 (0 =Not at all certain and 100 = completely certain).

3.3.1.2 Administration of the student questionnaire

The use of the questionnaire was vital because it helped to collect sufficient data in a relatively short time. A total of 1320 copies of student questionnaire were distributed to students in second and third years of lower level secondary schools i.e. approximately 560 copies in 28 second year classes (F2) and 760 copies distributed in 38 third year classes (F3). Sufficient introductory instructions were also written on the questionnaire. The administration of the questionnaire took place in classrooms during term time and during normal class time to maximize the questionnaire return rate. In many cases, I was able to meet the students in their respective classrooms during term time and most students were accessible and willing to participate. No questionnaire copy was sent to respondents via mail/email and no intermediary agents were used to reach out to the respondents. However, although the distribution and administration of the questionnaire was done personally, teachers and other school officials were in some cases contacted to help with the selection of participants and the distribution of copies of the questionnaire. These were first provided with clear information about the questionnaire and clear instructions about the survey so that they could take the work seriously (Bernhardt, 2013).

In most cases, time was reserved for question and answer sessions where I met face to face with the respondents to help those who had particular questions and or concerns regarding the survey. Oral clarifications were thus provided in person by the researcher directly to the respondents or by a proxy of a responsible staff at the school usually the dean of studies or the dean in charge of discipline. These clear and enough explanations were provided to help the respondents understand what they were required to do, why and how to proceed. This also helped build good rapport between me as a researcher and the respondents. Good rapport is essential to remove any feelings of threats or anxiety among respondents especially when dealing with children (Hatch, 1995). This was particularly relevant to this study as the respondents were adolescents. All the completed copies of the questionnaire were collected in person by the researcher. The completed student questionnaire copies were submitted to the school officials (teacher or head teacher) and collected by the researcher on a designated date.

Personal distribution and collection of the questionnaire by the researcher contributed significantly to the high return rate.

3.3.1.3 Structure of the questionnaire for teachers

The questionnaire for teachers was designed to gather data about the teachers' classroom assessment practices, i.e. about the forms of assessment they used, the purpose for which those classroom assessment forms were used, the timing and the content focus of assessment as well as the source of assessment tasks. The teachers' questionnaire was also used to gather information on the teachers' methods of providing feedback to students about their performance in classroom assessment tasks. The teachers' questionnaire was made of four parts covering five question themes (See Appendix 5).

Table 3.2: Structure of the questionnaire for teachers

Part	Theme and content	Number of Items	Scale
I	1. Demographic information	5	-
	<i>-Age, gender, teaching experience, professional qualification, in-service training</i>		
II	2. Assessment routine in a term time	24	4 -point Likert scale
	<i>-Focus of assessment, time of assessment and methods of providing feedback</i>		
III	3. Methods and Forms of classroom assessment	25	
	4. Origin of and content focus of assessment tasks		
IV	5. Purpose of assessment	18	
	<i>Assessment as learning, assessment of learning and assessment for learning</i>		
Total items		72	

As presented in Table 3.2, Part one aimed to collect each teacher's personal information about age, gender, teaching experience, as well as the teacher's professional qualification and in-service training level. All items on teacher's demographic data were formulated to collect information about individual respondents. However, these question items were designed in

such a way that anonymity was maintained and that no obtained information could be directly and easily linked to the respondent. According to Sudman and Bradburn (as cited in Cohen et al., 2007), it can be threatening and off putting when the provided information can be easily identifiable. It is essential that the respondents feel that the researcher is not seeking to collect their private information. Personal information was used for the recording and tabulation of collected data on the teachers' assessment practices. Demographic information was needed for establishing each respondent's context and explaining the teachers' differences in relation to their use of assessment forms.

Part two and three of teacher's questionnaire investigated the teachers' assessment routines and preferences in terms of assessment content and purposes. On the whole, part two sought to collect information that would help identify the content which was frequently assessed by teachers. Teachers were asked to indicate the focus of their assessment in terms of English language skills (i.e. writing, speaking, reading and listening) and other language features including grammar, vocabulary and pronunciation. Using a 4 point Likert scale where Never=1, Rarely=2, Sometimes=3, and Usually=4, Part two also investigated the extent to which teachers assessed non-achievement factors such as student's class attendance and effort. The assessment of non-achievement factors has been viewed by some researchers as problematic and ineffective in supporting learning due to the questionable reliability of used measurements and the concern that teachers may rely on their biased perceptions (Brookhart, 1994; McMillan, Myran and Workman, 2002). However, some research studies on teachers' classroom assessment practices have indicated that some teachers attach a great importance to non-achievement factors for which they assign grades to encourage their development (Alkharusi et al., 2014; Milnes and Cheng, 2008;). Although the role of non achievement assessment to learning has not been widely researched and confirmed, the continued teacher assessment of these factors implies that standard methods and measurements are needed to ensure that their assessment is done effectively.

In Part three, the same 4-point Likert scale was used to collect information on the forms of assessment that were frequently used in classroom. All participating teachers had to indicate their level of use of the stated forms of assessment on a 4-point Likert scale. This subscale was intended to measure the extent to which teachers used different forms of assessment in their

classrooms, the kinds of feedback that teachers provided to students and the source and the content focus of the assessment tasks. Part four of the questionnaire contained item questions that were categorized in three main themes related to assessment purposes namely *assessment as learning*, *assessment of learning* and *assessment for learning*. *Assessment for instruction* and *assessment to inform* were considered as being part of assessment for learning and assessment of learning respectively.

3.3.1.4 Administration of questionnaire for teachers

The teacher questionnaire was distributed in person by the researcher at all surveyed schools. A total of sixty six copies were distributed in forty one schools. As was done for student questionnaire, the teachers met in person with the researcher in order to provide them with sufficient information about the study and answer their questions. It was also an opportunity to discuss with them how the completed copies of the questionnaire would be collected. A date was usually agreed upon immediately regarding the venue and the day of the questionnaire collection. The filled in copies of the teacher questionnaire were collected from the surveyed schools. In some cases, teachers explained that they were busy with their teaching obligations and asked for extension for filling in the questionnaire. As a result, the teacher questionnaire return rate was 100% as the researcher was able to collect all sixty six copies of the teacher questionnaire.

3.3.2 Teacher interview

The interview as a tool of data collection is widely used in qualitative research methods. The increase in the use of the interview has followed the development of the constructivist perspective regarding knowledge as a product of human interaction which generates “*data* rather than *capta*” (Laing as cited in Cohen et al., 2007, p. 349). Laing argues that the interview is inter-subjective i.e. it allows the exchange of views between two or more people who share how they view and interpret the world they live in. In educational context, research interviews are defined as “communications that aim to consult teachers and students about their points of view, interpretations and meanings to help understand classroom dynamics” (Wilson and Fox, 2013, p. 116). These interview communications may be in form of interaction between the

interviewer and individual interviewees or between the interviewer and a group of interviewees (focus groups).

Interviews are also often classified in three categories depending on the level of formality. It can be classified as *unstructured* i.e. when conducted without a preset agenda and interviewer exercising little or no control, *semi-structured* i.e. where the interviewer knows the general themes of the interaction but does not preset questions, or *structured* i.e. where the interviewer preset all interview questions and determines their order (Nunan and Bailey, 2009). In mixed methods research, interviews can be used for collecting further information details to supplement data collected by use of other methods such as questionnaire and observation.

In this study, the term ‘interview’ is used to refer to the oral face-to-face interaction between the researcher and the teacher or a group of students. Individual interviews with teachers were conducted to help me inquire further on the respondents’ understanding of the statements in the questionnaire and how they justified their answers about their assessment practices. Semi-structured interviews were used where questions were generated from a set of themes that were preselected to guide the interview (Cousin, 2009). Some questions were also modified or adapted during interviews depending on the interviewees’ answers. In addition, follow up questions were asked to gather further in-depth information by probing some important responses that were provided by informants during the interview (Lodico, Spaulding and Voegtle, 2010; Walsh, 2001). Thus, a 15 to 25 minute interview was conducted to elicit more revealing information from each of the seven participating teachers whose classes were observed. This was also aimed at providing teachers with the opportunity to express their views in an interactive way and elaborate on the information that they provided in the questionnaire. In some cases, interview themes were modified after the review of the teachers’ responses in the questionnaire to address the topics or issues raised in the some teachers’ questionnaire. In other words, some questions were asked to follow up on the teachers’ answers about forms of classroom assessment and the extent to which they use such forms.

During the interview, open ended questions were used taking into account the use of ‘prompts’ i.e. the interviewer’s remarks aimed at clarifying topics or questions, and ‘probes’ i.e. interviewer’s questions that ask the respondents to elaborate on their answers (Cohen et al., 2007, p. 97). In this study, teacher’s interviews aimed to seek for clarification in relation to

what teachers did in classroom assessment especially why they were doing it. Cohen et al. argue that the “why” questions are very important in interviews as they help elicit more thoughts from the respondents. I used an interview schedule to pre-set some open ended questions that included the “why” items that were used as a guide during my interaction with interviewees. It has been argued that the use of an interview schedule can be essential for keeping the interview more focused on the topic for data relevance and ensuring that time and resources available for the research are used appropriately (Verma and Mallick, 1999).

Given the fact that interviews were held during term time, the interview schedule was set up to help save time of teachers who had a busy teaching load. I aimed to keep the interview brief but smooth and concise by focusing on issues that were not thoroughly covered by the questionnaire and observation. By using semi-structured interviews, I also had the freedom to ask more questions in the way that I found suitable for full investigation of the topics. The interview schedule listed the main and provisional probing questions that were to be asked to the teachers regarding the forms of assessment that they used to assess students and the methods that they used to provide feedback to their students. In addition, the pre-selection of topics and preparation of questions helped me to conduct the interview in more logical way, with questions sequenced following certain logic to facilitate easy understanding of the interviewee (Drever, 2003).

3.3.3 Student focus groups

The focus-group sessions were used to gather information from four groups of students at three schools. A focus-group is defined as “A small group made up of perhaps six to ten individuals with certain common features or characteristics, with whom a discussion can be focused into a given issue or topic” (Wellington, 2000, p. 241). The use of focus group also enables the researcher to gather collective views of a group of respondents who interact with each other which facilitates the emergence of new insights (Cohen et al., 2007). Focus groups also help the researcher to capture other useful information such as the level of consensus among participants and their disagreement.

The initial plan was to interview students in four groups, two groups from the city of Kigali and two from rural areas of the Southern Province. However, my invitations to four schools in Kigali to participate in the focus-group process were all declined citing reasons of schools busy schedules. For this reason, all four Focus-groups used in this study were from schools in the Southern province. A semi-structured schedule was used where a set of predesigned questions were asked with some follow up questions to seek clarifications from the participants' answers (See appendix 2a). At the beginning of every focus group discussion, the researcher explained the purpose of the study and provided more clarifications on the ethical issues related to their involvement in the study and the measures in place to protect them as respondents. They were informed that the whole discussion was audio recorded to facilitate later analysis of data.

The interview focused on three themes, namely the students' perception of assessment tasks, the methods that teachers used to assess and to provide feedback, and the students perceived levels of self-efficacy for speaking, reading, writing and listening to English. The researcher asked the participants in focus groups to provide their comments on all the seven key questions on the focus group schedule. To encourage every participant to get actively involved in the discussion, participants in each group were asked to take turns and answer and/or comment on the researcher's questions. In some cases, the researcher had to single out shy students so that everyone in the group could voice their opinion. Although some participants in each focus group were confident and always eager to give their comments, the overall observation was that participants in the focus groups were shy and reticent. As a result, the duration of the focus groups was shorter than was expected (see duration in Table 3.8). However, the students' focus groups generated rich discussions that revealed detailed students' beliefs, feelings and attitudes towards classroom assessment. Information from focus group supplemented data that was collected during the questionnaire survey.

3.3.4 Classroom observation

Classroom observation was used as another instrument of data collection in this study. On the one hand, observation was used to gather information on what happens in teachers' classroom assessment. Given the complexity of the classroom as a social setting where many things

happen at the same time, I decided to use a videotaping procedure to ensure that every aspect of what happened in the classroom was entirely captured. However, this was an ‘indirect observation’ i.e. observation conducted in the absence of the researcher (Curtis, Murphy and Shields, 2013). The indirect observation helps to avoid the unwanted effects that the presence of the researcher might cause to the research site. The use of video recorders in classroom research is thus useful not only because it helps avoid such effects but also because it catches a considerable amount of aspects of classroom interactions such as non verbal communication (Cohen et al., 2007). Video-recording is also advantageous in that the recorded data can be replayed several times after the recording process and the obtained information can be discussed later with different participants (Wilson and Fox, 2013). In this study, observation was expected to provide primary data that would offer a clear and accurate picture of assessment practices of the seven teachers who were expected to be video recorded while teaching their normal classes.

Some changes were made to my initial observation plans as a result of some schools’ lack of enthusiasm to participate in the classroom observation process. Only one out of three schools in Kigali City accepted my invitation to participate in classroom observation. However, no classroom observation was completed in this schools for reasons that were not always clear. In some schools, reasons such as busy teacher workload or an unexpected public event were given in the end resulting in the cancelation of observation. All four classroom observations that were completed were conducted in the Southern Province where five out of seven invited schools accepted my invitation to participate. In every observed classroom, a video-camera was set up in an appropriately selected place for capturing all teachers’ instructional and assessment practices.

The teacher-student interactional aspects of classroom assessment were recorded and transcribed using the Communicative Orientation of Language Teaching (COLT) scheme. COLT helps the researcher to record quantifiable data which in turn helps to make comparisons of what happens in different classrooms (Nunan and Bailey, 2009). The COLT scheme helped me to record the type of oral assessment interaction between teacher and students and their frequency of occurrence (see appendix 4a & 4b). Data on the type of assessment forms and teacher feedback was transcribed from the videotapes using the adapted Classroom Assessment

Practice Inventory by Zhang and Burry-Stock (2003). The frequency of their occurrences was recorded and statistically analysed. The recording of each teaching session lasted between a minimum of 30 minutes to the maximum of 50 minutes which equated one single academic hour. A total of 6 hours of teaching sessions were recorded. This is less observation time than what was expected due to reasons that I have cited above regarding schools' reluctance to be involved in the research. The initial plan was to observe ten teachers for 2 hours each which would give 20 hours of classroom observation. Although the observation time was reduced to six hours, the obtained data was believed to be highly informative and would be used to supplement other data sets.

3.4 Quality issues: Validity and reliability of research instruments

Necessary measures were taken to ensure that the research instruments used to collect data in this study were accurately designed in accordance with the existing theoretical and implementation principles of data collection instruments in order to avoid flawed measurements that may lead to biased results (Gideon, 2012). The recommended guidelines for the design of the questionnaire, interview, focus group and classroom observation were used to formulate question items and observation schedules that elicited the most valid and reliable information. The validity of information can be achieved in many ways. In qualitative research, these may include "honesty, depth, richness and scope of the data achieved, the participants approached, the extent of triangulation and the disinterestedness or objectivity of the researcher" and for quantitative research through "careful sampling, appropriate instrumentation and appropriate statistical treatment of the data" (Cohen, Manion and Morrison, 2011, p. 179). On the other hand, reliability of instruments and data can be enhanced by "controlling the conditions under which the data collection and measurement take place, training the researcher in order to insure consistency (inter-rater reliability) and widening the number of items on a particular topic" (Cohen et al. 2011 p. 201). In the current study, the researcher sought to maximise the quality of instrument and data by targeting a large number of respondents, piloting and using four different instruments to gather enough and rich data,

applying appropriate data analysis methods and using triangulation methods to verify accuracy of information.

The questionnaire was the main data collection instrument that helped to gather information from students and teachers. Its design was therefore given much consideration to ensure that valid and reliable information was collected. The questionnaires (see appendix 1 and 5) used multiple scales to collect measurable information on different aspect of the classroom assessment and students' perceptions. In both the teachers' and students' questionnaires, a 4-point scale was preferably used given its advantages to a 5-point scale. In their test of the reliability of a 4-point scale, Barton et al. (2008) found that a four point scale had a higher reliability level (Cronbach's Alpha) compared to a 5-point scale. The Cronbach's Alfa for this subscale was calculated at .62 with an inter-item correlation mean of $r=.18$, which are both marginally acceptable for a fewer item scale (Field, 2009). Another advantage of using a 4-point Likert scale over a 5-point Likert scale is that it limits the respondents' tendency to select the 'undecided' choice. Instead, it compels the respondents to indicate their tendencies on one of the four choices (Baker and Ellice, 2011). In this study, a 4-point Likert scale was used in lieu of the 5-point Likert scale to minimize the non decided cases of the respondents.

Part three of the student questionnaire used an 11-point scale (see Appendix 1a). The scale was adopted from Bandura (2006) who argues that an 11-point scale is more reliable than a scale based on fewer points. He stresses that more points provide options and alternatives for respondents and thus effectively capture differences among respondents. A similar 11-point Likert scale was used by Chen (2007) in his study on the relationship between the listening self-efficacy beliefs of Taiwanese EFL college students and their English listening proficiency. The scale was proven to have internal consistency and its reliability was calculated above $\alpha=.90$ Cronbach's alpha coefficient. Each subscale used to measure the students' self-efficacy in each language skill had a Cronbach's alpha above $\alpha=.90$. It was $\alpha=.91$ for the subscale on speaking, $\alpha=.93$ for the subscale on writing, $\alpha=.92$ for reading and $\alpha=.94$ for listening. The whole self-efficacy scale had $\alpha=.97$ of internal consistency as a measurement of the concept of self-efficacy.

It is also important to note that while Bandura's (2006) 11-point scale was maintained, the item statements used in this study were adapted from the English language curriculum issued by the Rwandan ministry of education in 2011 and still in use in lower secondary schools during the time of data collection. Reference was made to the national English language curriculum in order to construct the relevant scale that included question items reflecting the learning content. For example, the statements on sports and leisure in Bandura's (2006, p. 327) scale on leisure asks students to rate the degree of their confidence by recording a number from 0 to 100 on how certain they are that they can "Learn the skills needed for team sports (for example, basketball, volleyball, swimming, football, soccer)". This was matched to the topic of 'sport and leisure' in the curriculum content with a subtopic of "talking about holidays". The learning outcome stated in the curriculum was that "by the end of this sub-topic, learners will be able to talk about what they did in the holidays". From this statement of outcome, a corresponding statement about the learners' self-efficacy for speaking was formulated as "I can give a talk about how I spent the last holidays" (See statement 3 in part III of appendix 1a). Other statements about reading and listening tasks as well as writing performance activities outlined in the curriculum were included in the questionnaire. The respondents had to indicate a specific level at which they believed they could perform each of the stated tasks. They were asked to indicate the extent to which they agreed with the provided statements about their perceived capabilities to successfully perform the tasks related to learning outcomes as outlined in the curriculum.

The quality of the questionnaire was also enhanced by producing a copy of the questionnaire in the native language of the respondents. Although the questionnaire for students was initially written in English, I requested my fellow research students and other experienced researchers to review the English version of the questionnaire for both grammatical and technical errors that could distort the intended meaning (Krysik and Finn, 2013). I proceeded with translation of the questionnaire from English to Kinyarwanda, which was the native language of the respondents. Translation of the questionnaire into the respondents' mother tongue is recommended to facilitate easy and full understanding of question items and to raise the degree of information accuracy and reliability. The translation of the questionnaire also needs to focus on meaning rather than on literal translation (Brace, 2013). During the

translation of the questionnaire, high consideration was given to producing the Kinyarwanda version that was culturally meaningful and easy to comprehend. Bilingual researchers, my fellow compatriot teachers and experienced translators were contacted and asked to validate the translated copy of English to Kinyarwanda version of the questionnaire.

The listed names of assessment forms and tasks in part three of the teacher questionnaire (see Appendix 5) are based on the Assessment Practice Inventory by Zhang and Burry-Stock (2003) which lists 67 items that reflect the range of assessment activities including test construction, scores interpretation, grading and communication and use of assessment results. In this study, focus was put on items relating to assessment design, assessment administration, grading and communication of assessment results. In addition, the Classroom Assessment Preference Survey Questionnaire for Language Teachers (CAPSQ-LT) developed by Gonzales and Aliponga (2011) was adapted and used to gather data on the purpose of teachers' classroom assessment in part four of the questionnaire. The CAPSQ-LT was adapted because no relevant and appropriate instrument was available for investigating the research classroom assessment practices. Gonzales and Aliponga used their scale to collect data on the assessment purposes of secondary school Japanese and English language teachers in Japan and Philippines. Good psychometric properties of the scale were confirmed by the high score of the Cronbach's alpha coefficients that ranged between .82 (for assessing to inform) and .93 (for assessment as learning). However, as some changes were applied to this questionnaire after a pilot survey (see section 3.6), its' internal consistency was reduced as a result of the reduction in number of question items. The internal consistency levels became $\alpha=.73$ for items on assessment as learning (item 1-4), $\alpha=.60$ for assessment of learning (items 5-8 & 16-18) and $\alpha=.74$ for items on assessment for learning (item 9-15). According to Field (2009), the level of consistency from $\alpha=.60$ is acceptable for a fewer item scale. However, the total index for the whole subscale on assessment purpose (18 items on assessment as learning, of learning and for learning) was $\alpha=.83$ with inter-item correlation of $r=2.2$.

Other methods that were used to ensure that quality of instruments data was maintained included the piloting of research instruments (see section 3.6 on the pilot process). The pilot process helped to foresee and troubleshoot some problems that could arise during the main questionnaire survey and during interviews, during focus groups and classroom observation. In

addition, the collection of data from different respondents (students and teachers) by use of different instruments allowed for data triangulation which is one of the methods of enhancing the reliability of data. Information provided by the students was intended to help compare and verify the consistency of the teachers' answers on how often they used different forms of assessment in their classrooms. This was also compared to the information gathered from the classroom observation and the teachers' interviews as well as focus groups to identify and provide explanations to any inconsistency.

3.5 Research participants

Questions on the respondents' background were formulated to gather personal information about each participating student and teacher. This information was collected for descriptive purposes and to facilitate analysis of data for each research question in the context of relevant sub-groups and to understand differences that existed within different group clusters of informants. It also helped to gather small but important details essential for a holistic understanding and interpretation of the study results. The participants in this study were drawn from a population of teachers and students from private, public and government aided schools in Rwanda.

The study targeted the lower level of secondary schools where the national English language curriculum published by MINEDUC in 2010 was used for the teaching and learning of English. A study population is defined as the total number of all elements under investigation from which a sample is drawn and upon which the results from the studied sample are applied (Babbie, 2013; Bless, Hugson-Smith and Kagee, 2007; Hartas, 2010). A sample of informants were selected to participate in this study and to represent the population that could not be studied entirely due to time constraint, difficulty accessing the entire population, financial costs involved and data manageability among other factors (Bailey, 2008; Cohen et al. 2007; Bless et al., 2007). The participants were all students and teachers from selected schools who were involved in the process of information gathering as respondents through questionnaire, interviews, observation and focus groups.

3.5.1 Sample size and sampling process

Different sampling procedures were used to select the appropriate number of participants necessary for a good representation of the target population. According to Cargan (2007), the main objective of sampling is to draw a sample that is representative enough to generate generalizable results. To increase the representativeness of the sample, appropriate measures were taken to involve informants who had the distinctive features that matched the characteristics of the target population. A sample is representative when “its aggregate characteristics closely approximate those same characteristics in the population” (Rubin and Babbie, 2009, p. 134). In order to reflect the reality of the English language teaching at the lower level secondary schools in Rwanda, I used stratified sampling technique to select participants from different subgroups of the population. By using this technique, the selected respondents were expected to be drawn from each subset of the population strata (Bailey, 2008; Hall, 2008; JHA, 2014). The strata in this study were at the school location level (rural vs urban), the school learning mode (Day vs Boarding) and the school status (Private, public, or government aided).

Data was collected from the schools located in rural sectors of the Southern Province districts and from schools located in the capital city of Kigali, the cities of Huye, Nyanza and Muhanga (See Table 3.3). This meant that the sample of this study was comprised of students and teachers from urban schools in Kigali city and from urban and rural sectors of five districts in the Southern province. Both boarding and day schools, private and public schools as well as government aided schools were all targeted. The teachers were of all ages, of a diversified range of experiences and various professional qualifications. The students had to be in their second or third year and it was assumed that they had participated in different classroom assessment tasks.

In this study, it was not possible for the researcher to determine the exact size of the population. In other words, the exact total number of all students enrolled in F2 and F3 in the Southern province and Kigali city was not available at the time of data collection. Therefore, on the one hand, the non-probability sampling technique was used to estimate the overall size of the sample and the suggested guidelines on non-statistical methods of sampling were applied

whereby any sample between 200 and 1,500 participants is suitable for a population size of over 400 (Johnnie, 2012). On the other hand, the size of the sample was determined after consideration of the smallest number of the respondents in the smallest subset (Borg and Gall, 1979). The number of respondents in the classroom, which was the smallest subset, was estimated at 35 students as per classroom-student average ratio in 2015 (MINEDUC, 2015b). It was therefore decided to select at least 20 students in each classroom which was expected to produce a total sample of 1320 students and 66 teachers. 1258 student questionnaire were returned, i.e. 95.3% return rate, with 100% return rate for teacher's questionnaire. According to Cohen et al. (2007), a sample size of 1,060 is appropriate for a population size of up to 150,000. Although the population size of this study was not identified, the sample size of 1,324 (teachers and students) was judged appropriate and would be applicable even with probability sampling, given that all the students in the lower level of the secondary schools in the Southern province and Kigali city were estimated at 123,007 in 2015 (MINEDUC,2015).

To select the participants at the province, district and schools levels, I used the non-probability sampling techniques where not all elements of the population had equal chance of being selected. Cohen et al. (2007) explain that researchers using non-probability sampling select the sample with full knowledge that it does represent itself not the entire population. The researcher simply targets and selects sample elements that meet certain preset criteria (Clack and Adler, 2014; Marlow, 2010). Therefore, by using non-probability sampling technique, not all provinces, districts or all secondary schools in all provinces of Rwanda had equal chance of being selected and neither all students nor all teachers in 2nd form (F2) and 3rd form (F3) of lower level of secondary schools in the country had equal chance to participate. The investigation targeted teachers and students of F2 and F3 of the lower level secondary schools located in rural and urban areas of the Southern Province and the schools located in the capital city of Kigali where English was taught as a main subject. As of 2015, the Southern Province had the highest number of secondary school students in the country. It also had a total of 85,216 students enrolled in F1, F2 and F3, which was 25% of all students in Rwandan lower level of secondary schools and the highest of all four provinces and Kigali City (MINEDUC, 2015b).

As illustrated in Table 3.3, a total of 1258 students and 66 teachers were selected from 41 schools located in eight districts of Southern province and Kigali city. Seven of these teachers from five schools were also observed and interviewed (see section 3.5.3) and four student focus group interviews were carried out at three schools (see section 3.5.4). These sample sizes were chosen in consideration of the nature of the current study, time and resources available for the study and the need for the maximisation of accuracy (Sarantakos, 2013).

Table 3.3: Sample size by province

Province	District	Number of participating schools	Number of Participants	
			Students	Teachers
Kigali City	Nyarugenge	3	98	5
	Kicukiro	3	77	4
Southern Province	Huye	11	377	20
	Gisagara	4	114	6
	Nyanza	9	276	14
	Ruhango	4	156	8
	Muhanga	7	160	9
Total	7	41	1258	66

Since some secondary schools did not have the lower level, I used the purposive sampling technique to guide the process of selecting the participating schools. This technique allows the researcher to control the variables and to form the right sample by selecting participants whose characteristics are relevant to the study (Bryman, 2012; Sharma, 1997). Therefore, I used this technique to select schools that had F2 and/or F3 of lower level secondary school and where English was taught as a subject. In addition, the same technique was used to select schools where sample classes were drawn. The schools were selected from both rural and urban areas in order to control potential influential factors such as school's environment or factors related to the socioeconomic status of participants.

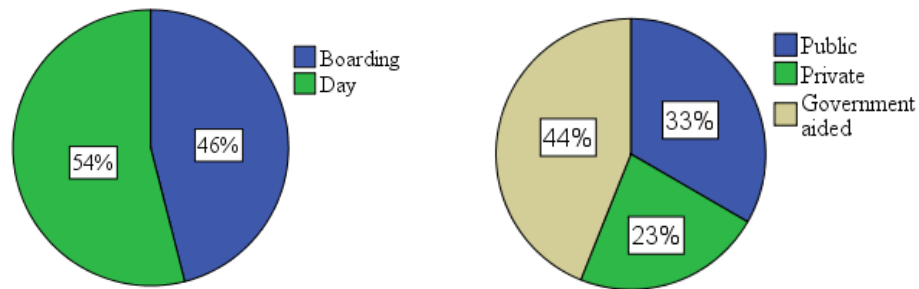


Figure 3.2: School Learning mode (n=41) Figure 3.3: School Status (n=41)

In addition to selecting participants from private, public and government aided schools (see Figures 3.3), participants were also selected on basis of their school learning mode (see Figure 3.2), i.e. depending on whether they studied in boarding schools or day schools. In Rwanda, the lower secondary schools are classified in two categories. The first category is comprised of Boarding schools where students who obtain higher grades in the National Primary School Leaving Examination are usually enrolled. The second category consists of day Schools which usually received students with lower grades who did not satisfy entry requirements of the boarding schools (see Section 1.2). Boarding schools are usually public or Government aided- schools which also often have upper level of secondary education. As indicated in Figures 3.2 and 3.3, up to 54% of the selected schools were day schools and 46% were Boarding schools. The highest number of surveyed schools were also government aided (44%).

Some studies have argued that well qualified and more experienced teachers often prefer to work in schools located in urban areas (Berryman, 2000; Mulkeen, 2006). This can have significant impact on learning as some studies have linked aspects of teacher qualification and experience to the students' achievement (Clotfelter, Ladd and Vigdor, 2006; Ma, 2008). Similarly, students' achievement has been found to be related to the students' self-efficacy through what is commonly known as 'experience of mastery' (Bandura, 1997). Students' self-efficacy is likely to increase when students experience high achievement in their learning. Hence, it can be expected that students in urban schools where many teachers are assumed to be experienced and well qualified, would report high levels of self-efficacy compared to their counterparts in rural areas. For this reason, the rural-urban variability factor was taken into

consideration. Relevant data was collected to investigate differences in the students' self-efficacy based on the location of their schools.

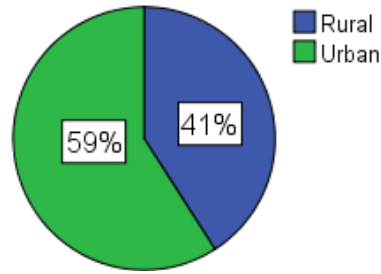


Figure 3.4: Location of the surveyed schools (n=41)

Schools that were located in rural sectors or were at ten kilometres from the city centre were considered as rural. A school that was located within less than ten kilometres from the city centre or in a urban areas of a district was considered as urban. This distance range was considered realistic because cities in the Southern province are generally small and rarely cover a ten square kilometre area. All schools in Kigali capital city were considered urban. As can be seen in Figure 3.4, more than a half of the surveyed schools were selected from urban areas (59%) while schools in rural areas represented 41%.

Participation of both male and female students in this study was a preset requirement. The results of analysis of the demographic data showed that both male and female students participated at the same level as anticipated. Out of 1258 students who participated in the study, 50% identified themselves as male and 49.9% as female while 0.1% of the respondents did not specify their gender. Their age range was between 12 and 16 with some students indicating that they were above 16. In some cases, students were in their 20s particularly in day schools. Figure 3.5 indicates the age range per gender of 1242 students who indicated their age.

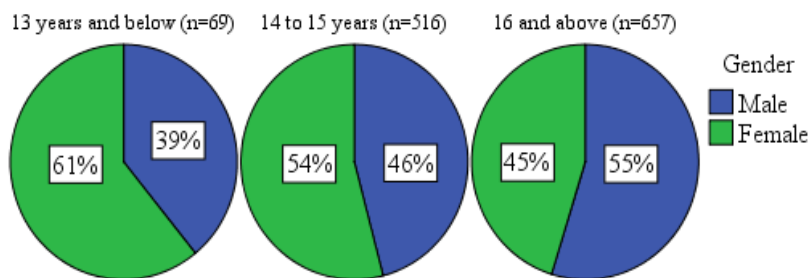


Figure 3.5: Students' age by gender (n=1242)

As displayed in Figure 3.5, a small number of the respondents were aged 13 and below with the majority of them being female (61%). The figure also indicates that while the number of male respondents increased as the group age increased (from 39% for 13 years and below to 55 % for 16 years and above), female respondents decreased as the group age increased (61 % for 13 years and below to 45% for 16 years and above). There was a slight difference in age distribution according to school location. Students in rural school were younger than students in urban schools where up to 55% were 16 years old or above against 50% in rural schools. The age difference was also observed across the two class levels with F3 having more students aged 16 and above while F2 had younger students mainly below 15. These differences in age groups were expected given that the secondary school age in Rwanda was between 13 and 18 at the time of data collection (EPDC, 2014). Thus the majority of students in F2 were expected to be 13 or above i.e. having completed their 6 year primary education. Also, as was expected, the number of students aged 16 and more increased with class levels, twice higher in F3 than in F2.

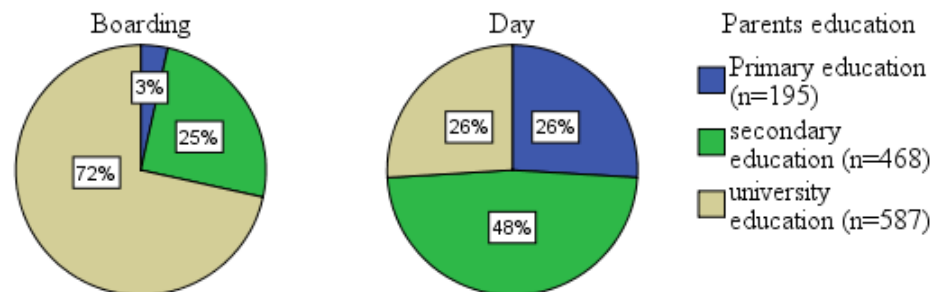


Figure 3.6: Students' parents education by school learning mode

Information was also gathered regarding the level of education of the respondents' parents. This was collected to control factors related to the students' background such as such as parental influence exerted in the form of aspirations and persuasions. Parental influence has been found to be important in the formation of children's self beliefs (Schunk, 1995). As displayed in Figure 3.6, up to 72% of students who studied in boarding schools had parents who had university education. On the other hand, only 3% of students in boarding schools had parents with primary school education. Figure 3.6 also shows that the majority of students in day schools (48%) had parents with secondary education.

Data was also collected about teachers' age and gender, qualifications, their experience and their in-service training. The results show that 68% of teachers who participated in the study were male and 32% were female. The majority of teachers had bachelor's degree (91%) and some had secondary school certificates (3%) and masters' degree (6%).

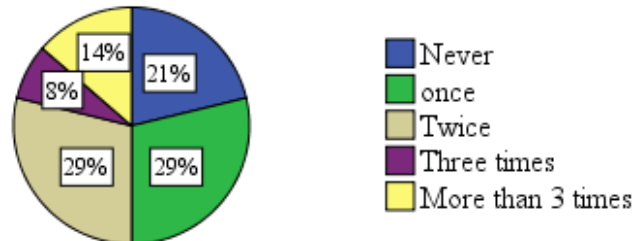


Figure 3.7: Teacher in-service training in a 5 year period (n=66)

Most teachers were aged between 25 and 35, representing 55% of all teachers who participated in the survey. As presented in Figure 3.7, half of the teachers reported having had one or no training at all in a five year period. These included teachers who indicated that they never had any training in assessment (21%) and others who said that they had at least one training session within the same period (29%). Only 14% of all teachers said that they had attended training in assessment more than three times in the same five year period.

3.5.2 Participants in the questionnaire survey

Both teachers and students participated in the data gathering process by providing their answers to research question items in the questionnaire. As was indicated in Table 3.3, a sample of 1258 students and 66 teachers took part in the questionnaire survey. Up to 40 copies of the student's questionnaire and two copies of teacher's questionnaire were distributed in each surveyed school. Given the fact that the average pupil-classroom ratio was 35:1 (MINEDUC, 2015b) it was intended that 20 copies of student questionnaire and 1 teacher questionnaire would be distributed in each classroom. All the participating schools, students and teachers were selected by use of a sampling framework that contained general information about the prospective participants such as names of schools, and their addresses (Denscombe, 2014; Saris and Gallhofer, 2014).

Table 3.4: Gender and class level of the students who participated in the survey

			Class level		Total
			F2	F3	
Gender	Male	n	278	351	629
		%	22%	28%	50%
	Female	n	267	361	628
		%	21%	29%	50%
Total		n	545	712	1257*
		%	43%	57%	100%

*The gender of one student was not indicated

The sampling framework was obtained from the sector officers in charge of education. It listed the names of public and private schools that had lower level of secondary education and were located in rural areas of Southern province and in cities of Kigali, Nyanza, Muhanga and Huye. The lists were also indicative of all the levels of secondary education available at each school (lower level, upper level or both). This information was used to select classes and teachers who were invited to participate in the investigation. As can be seen in Table 3.4, out of 1258 students who participated in the study, 43% were enrolled in F2 and 57% were in F3.

The secondary school age in Rwanda ranges between thirteen and eighteen years (World Bank, 2014). Although the majority of students of F2 and F3 of secondary schools who participated in this survey were generally between thirteen and sixteen years of age, some cases of students under twelve and above sixteen. Participants in the questionnaire survey were also from different socioeconomic background. The sample also included both males and females and both were considered equal participants throughout the study. However, as mentioned earlier, the student sample from each school was selected to represent the entire class population that could not be studied entirely given the high student-classroom ratio.

Teachers of English in the 2nd and 3rd forms of lower level secondary schools at selected public, government aided and private rural and urban schools were invited to participate regardless of their age, gender, teaching experience or qualification. Although preference was to see all these traits represented in the sample, it was not part of the eligibility criteria for participation in the survey. However, the participating teachers were requested to provide some personal information in order to control possible key variables such as age, teaching

experience, qualification and many others factors that may influence teachers' everyday teaching practices. Having the status as teachers of English in F2 and/or F3 automatically yielded the eligibility to teachers to take part in the survey upon signing of consent forms (See appendix 6).

Table 3.5: Gender and taught class of teachers who participated in the survey

		Gender		Total
		Male	Female	
Class Level	2 nd Form	18	10	28
	3 rd Form	27	11	38
Total		45	21	66

Similarly to the student survey, both male and female teachers participated in the study. Table 3.5 shows that fewer female teachers participated in the survey (42%) compared to their male colleagues. This was anticipated given the fact that the proportion of female teachers in Rwandan secondary schools is often lower than male teachers. As of the academic year 2016, female teachers in secondary schools represented 30.3% of the teaching staff (MINEDUC, 2016).

Table 3.6: Teachers' gender by school status

			Gender		Total
			Male	Female	
School status	Public	Count	14	7	21
		% within Schools status	66.7%	33.3%	100.0%
	Private	Count	13	2	15
		% within Schools status	86.7%	13.3%	100.0%
	Government aided	Count	18	12	30
		% within Schools status	60.0%	40.0%	100.0%
Total		Count	45	21	66
		% within School status	68.2%	31.8%	100.0%

Table 3.6 also shows that the majority of female teachers taught in public and government aided schools while only 13.8% taught in private schools. Necessary measures were taken and maintained to ensure that research both male and female teachers understood well what the purpose and objectives of the study were. It was also important that participating teachers developed good rapport and confidence in me as a researcher (Hitchcock and Hughes, 2002). Creating good understanding with teachers helps them feel free to provide information concerning their individual ways of assessing students. Some teachers could have declined to share information but establishing the relationship based on mutual trust with them made every participant feel free.

3.5.3 Participants in the teacher interview

A limited number of teachers of English in F2 and F3 of lower level at the selected secondary schools were invited to participate in individual meetings with me to discuss issues concerning their forms of classroom assessment, purpose of their assessment and feedback that they provided. All seven teachers who had been observed and participated in the questionnaire survey were invited to volunteer meeting with me regardless of their profiles. However, more attention was given to teachers who had provided more revealing information worth following up. A semi-structured interview schedule was used to gather more information on key themes surrounding classroom-based assessment practices. The interview was centred on three themes of assessment forms, purpose of assessment and teachers methods of providing feedback. The first part of the interview focused on my classroom observation remarks and aimed at finding out how teachers understood assessment and how and why they chose the assessment activities that they used. The second part inquired on the teachers' assessment routines mainly in terms of purposes, assessment tools and feedback that they provided to students.

Teachers' interviews were conducted in either English or Kinyarwanda and in some cases both languages were intermittently used in one interview depending on the preference of each interviewee. Some researchers argue that respondents may not feel comfortable discussing in a foreign language (Bailey, as cited in Cohen et al., 2007). Hence, given the fact that English is not the native language of Rwandan teachers, local teachers of English in Lower level were given the option to choose the language to use during interview between English and

Kinyarwanda. This was done to help teachers with low level of confidence in English and who might feel more comfortable in their native language.

Table 3.7: Description of the teacher’s interview

Teacher (initials of the Pseudonym)	Gender	Class level	Venue	Duration	Date	Language
PL (Phil)	M	F3	Classroom	15 min	02/06/2016	Kinyarwanda
FR (Fred)	M	F3	Teachers’ room	25 min.	02/06/2016	English
MR (Mary)	F	F2	Teachers’ room	16 min.	09/06/2016	Kiny. & Eng.
VN (Vincent)	M	F3	Teachers’ room	17 min.	09/06/2016	Kiny. & Eng.
JS (Jessica)	F	F2	Teachers’ room	15 min.	09/06/2016	Kiny. & Eng.
SF (Stephen)	M	F3	Teachers’ room	24 min.	09/06/2016	English
SL (Selah)	F	F3	Head teacher’s office	17 min.	15/06/2016	Kinyarwanda

Although it was assumed that teachers were capable of expressing themselves in English, the Kinyarwanda version of teacher’s questionnaire was made available for teachers who would prefer to use their mother tongue for interactional comfort.

The English version of the teacher’s interview schedule was particularly intended for teachers whose native language was not Kinyarwanda. When English language was officially adopted as the language of instruction for secondary and tertiary education in the country, the ministry of education started recruiting teachers of English from English speaking neighbouring countries to help train and mentor local Rwanda teachers. In some cases, these regional English language teachers were hired to teach English in secondary schools across the country. Many of these regional teachers could not speak Kinyarwanda and therefore needed to be interviewed in English. All questions in both versions were the same and audio recording instruments were used to record each interview upon consent of the respondents.

In the effort to maximize the respondents’ freedom, I conformed to every respondent’s choice of interview venue and time. Teachers were free to select the place where they wanted the interview to be held, either in their offices, in another school premises, or at any other place that they found convenient. This had to be decided before the beginning of each interview to ensure that no time was lost in search of venue. As can be seen in Table 3.7, most teachers’

interviews were held in the teachers' room. These were conducted during break time when no other teachers were using the room. However, in case where the teachers' room was unavailable, the school administration allowed us to use their offices. For example, my interview with Teacher SL was held in the head teacher's office. Only my interview with teacher PL was conducted in the classroom because other places were being used. The interview was conducted during lunch time and no students were in or around the classroom. Overall, both teachers and administrative staff in schools that participated in the study were all collaborative and supportive.

3.5.4 Participants in student focus groups

Each group was made up of five to eight students from the same class. Each group was comprised of male and female students from the observed classrooms. I first introduced my research to students before the classroom observation and invited them to participate in the interview on a voluntary basis. Of the four interviews that were conducted, two were held in the teachers' room and two others were conducted in one office of the dean of studies. English teachers were helpful in finding and freeing up calm venues that were appropriate for recording.

Table 3.8: Description of students' focus groups

Focus Group	Class level	Number of participants	Duration	Date of interview
A	F3	8	20 min	02/06/2016
B	F3	5	23 min	02/06/2016
C	F2	6	15 min.	09/06/2016
D	F3	6	16 min.	09/06/2016

Each focus group members were asked about the methods used by their teachers in everyday classroom assessment. As this was a semi-structured interview, I followed the order of the main questions that were prelisted on the interview schedule. However, they were also asked follow up questions to obtain further clarification on what was said. The interview also touched on the students' self-efficacy for using English in spoken and written communication and their perceptions of teacher classroom assessment in general.

3.6 Pilot process

Piloting the questionnaire helps to increase its levels of reliability, validity and practicality in terms of clarity, length, and attractiveness and removes redundancy among other features (Cohen et al. 2007). The piloting of a questionnaire aims to examine whether it measures what it is intended to measure, whether it adequately collects the needed information (Brace, 2013). In the current study, conducting pilot survey helped to identify and remove potential problems such as ambiguous or irrelevant questions that could negatively affect the main survey. Testing this instrument also helped to identify other errors which were then corrected before the process of gathering the main data.

After the local ethics clearance was obtained, a pilot was conducted in three schools of the southern province where three teachers and 75 students filled in the questionnaire. In addition to my own observations, I solicited the respondents, especially teachers to provide their comments on how they found their questionnaire and what they thought about the students'. Although the students' questionnaire appeared to be generally easy to understand, some issues were identified including that students were taking much time to fill in the questionnaire. Students were taking up to 40 minutes average to complete the questionnaire while it had been designed to take around 15 minutes on average. Following this observation, some items were removed, rearranged or modified to reduce the time load and to facilitate good comprehension of the questions. For example, the number of question items about the self-efficacy scale (see appendix 1a) was reduced from 28 in total to 20. Five statements were provided for each language skill instead of seven statements that were used in the first previous scale (See part III of the student questionnaire in appendix 1a). Also, in addition to the high time load, it appeared that some questions were confusing to some students. The questionnaire was thus revised and instruction was provided for every question item. The wording of some question items was also changed and examples provided to allow easy understanding. It is for this purpose that examples were added to the question item about the forms of assessment used in the classroom (see question 10 of the student questionnaire in appendix 1a). The students' questionnaire was compressed from 3 pages to 2 pages and the teacher's questionnaire from 6 pages to 4 pages. However, this was done carefully to preserve the objectives of the survey and to ensure that the original purpose of each subsection in the questionnaire was not altered.

Two pilot observations were also done to identify issues that could have effects on the observation data. One pilot observation was conducted in the UK at the University of York and another in Rwanda prior to the main survey. The UK pilot was conducted to complete a pre-survey view of how classroom observation would be carried out in the place of data collection. This was a 50 minute long observation carried out in a classroom of nearly 20 students. It provided useful insights and an opportunity to identify possible issues with the observation schemes. As a result of this observation, the scheme was modified so that it could capture all aspects of teacher-student interaction.

The modified scheme included oral questioning in addition to used forms of assessment. The second pilot observation was done to examine possible issues applicable to the Rwandan context where the survey was to be conducted. It was done in one classroom in the Sothern province where I requested the teacher to record his teaching session. After I demonstrated to him how to use the camera in the classroom for recording his teaching session, the teacher recorded himself teaching for a period of two hours. Some of key issues that were identified included that it was difficult to find a spot in the classroom where the camera could capture the entire class activities. The camera could capture only one part of the class, usually two thirds of the students. There were also problems of the use of camera where, despite my guidance to him, the teacher tended to move the camera constantly in order to capture the entire classroom. However, this affected the quality of the recorded data as it interrupted the teaching process and was seen as a distraction to students who turned back to see what the teacher was doing. Technical issues were also observed such as camera running out of power and internal memory running small. These were all noted and fixed during the main survey.

3.7 Data analysis procedures

Both qualitative and quantitative methods of data analysis were used in this study. The quantitative methods were used to analyse all data collected from the questionnaire survey which investigated the teachers' classroom assessment practices and the students' perceptions and self-efficacy beliefs. A part of data collected from classroom observations was also

analysed by use of statistical procedures. On the other hand, the qualitative methods were used for the analysis of data from teachers' and students' interviews.

3.7.1 Quantitative data analysis

Descriptive statistics was used to analyse data on research question one regarding the assessment practices that teachers of English used in their classrooms mainly regarding their use of assessment forms and how they provided feedback to their students. Both SPSS and Excel software were used for computational purposes, to calculate the frequencies, mode and median or to compute means and standard deviations of numerical data from the questionnaire survey and from the classroom observation. Although the majority of the data on research question one was obtained from the teachers' and students' questionnaire survey, some quantitative data was obtained from Classroom observation. This consisted of the number of observed occurrences of assessment forms used by teachers during their classroom interactions with students (see Table 4.3), as well as the recorded frequencies of use of feedback provided to students in the classroom. Means and Modes were also used to analyse data on the students' perception of their teachers' classroom assessment practices (See Table 4.10). Descriptive statistics was also used to explore data obtained on research question two specifically on its sub-research question 2.1 regarding the levels of the students' self-efficacy. The students' self-efficacy scores on the stated tasks were averaged to obtain the mean score representing every student's self-efficacy level in each language skill.

In addition, inferential statistical methods were used to explore the differences between groups of respondents and to measure the levels of relationship between different variables. As most data in the current study was ordinal, the non-parametric statistics was found more appropriate than the parametric tests. The latter could be used for the analysis of the continuous data. However, most of the continuous data obtained in this study lacked the necessary characteristics required for the use of parametric tests such as normality of data distribution. Therefore, the Mann Whitney U test was used to determine the statistical significance of the self-efficacy differences between two groups of students (eg. male vs female, day vs boarding schools). The statistical differences between more than two groups (e.g private vs public vs government aided schools) were measured using the Kruskal-Wallis Test which is appropriate

for analysis of differences based on ordinal and non-normally distributed data. Data on sub-research question 2.2 and 2.3 which sought to examine the relationship between variables relating to the teachers' classroom assessment practices and the students' self-efficacy was analysed by use of the Spearman rho correlation test. This is a non-parametric test based on data ranking and used to explore the relationship between ordinal variables. The correlation coefficients and the P values were reported for significance of relationship and the coefficient of determination was calculated from the obtained correlation coefficients to examine the level of variance shared between the tested variables.

3.7.2 Qualitative data analysis

The content based analysis was used to sort, arrange and classify textual information of non-numerical data from classroom observation, teacher interviews and students' focus groups. Specifically, the manual coding methods were used to analyse qualitative data obtained by use of three instruments namely the Classroom Observation Schedule (COS), the Students' Interview Schedule (SIS) and the Teachers' Interview Schedule (TIS) which were both used during face-to-face interactions with the respondents. The Classroom Observation Schedule (COS) was used to collect further information on the teachers' assessment practices and capture important qualitative information that helped to get full insights on what happened in the classroom. Every recorded video was analysed and notes were written highlighting all the assessment features observable in every lesson. These notes, together with data from teacher interviews and student focus groups were used to describe the classroom assessment routines with more details. The results from the analysis of qualitative data were also intended to be used for data triangulation purposes in order to cross check the accuracy of each data set. Although the quantitative data was large and led to very important results, the analysis of the qualitative data was vital because it helped to add more clarification to the results from the quantitative data.

Data from both students and teachers' interviews was analysed through manual coding i.e. without using nvivo software. This is a process that helped to identify key categories to which evidence from the collected raw data was linked. Coding is defined as a "process by which a text is examined thematically according to certain categories (codes) which are either

predetermined or emergent from the data” (Evans, 2013, p. 163). The main codes in the current research were predetermined and the coding process consisted of examining the extent to which these codes appeared in the respondents’ information. This data analysis was completed in three stages consisting of (1) initial preparation of data, (2) the first and second coding cycle followed by (3) the synthesis stage.

Interview transcripts excerpt	Codes
<p><i>Researcher:</i> Have you assessed in this week? Assessment can refer to tasks that you give to students and which you may or may not mark them.</p> <p><i>PL:</i> yes I have....,</p> <p><i>Researcher:</i> May I know what the kind of assessment you gave?</p> <p><i>PL:</i> <u>quizzes... and exercises</u></p> <p><i>Researcher:</i> Ok, and what was the purpose of your assessment?</p> <p><i>PL:</i> It was to see...if students... what students...after the lesson, what they gained from the lesson</p> <p><i>Researcher:</i> <u>What they gained</u></p> <p><i>PL:</i> Yes</p> <p><i>Researcher:</i> Ah... did you use rubrics? Rubrics I mean...can be a separate document that provides more details about the assessment...</p> <p><i>PL:</i> I use <u>rubrics when it is a test</u> out of 100 points. But if it is out of 10 or 15 points I don’t use it</p> <p><i>Researcher:</i> That test out of 100, when is it given, is it after...</p> <p><i>PL:</i> This is <u>given at the end</u>, after..., it’s done once in a term</p> <p><i>Researcher:</i> Ok so when you use it, do you share it with students?</p> <p><i>PL:</i> it is <u>given through instructions</u> on the test paper, how they will do the test..., the number of questions etc</p> <p><i>Researcher:</i> ok now about the methods you use to assess. Which ones do you normally use?</p> <p><i>PL:</i> Ok you can use...we use ...<u>question and answer</u>: you ask and one student answers, you can... for example I’ve just used <u>pair work</u>, in pairs of two using one <u>answer paper</u>...yes... both written and <u>oral</u> are used</p> <p><i>Researcher:</i> What do you mean by oral?</p> <p><i>PL:</i> <u>asking students</u> in the classroom and they answer</p>	<p>“quizzes... and exercises”</p> <p>“what they gained”</p> <p>“rubrics when it is a test”</p> <p>“given at the end”</p> <p>“given through instructions”</p> <p>“question and answer”</p> <p>“Pair work”</p> <p>“answer paper” “written”</p> <p>“oral”</p> <p>“Asking students”</p>

Figure 3.8: Example of first cycle analysis of teacher interview using manual coding (Transcript extract from interview with Teacher PL)

During the initial preparation, the transcribed interview data was reviewed for typing errors or incorrect translation of the respondents’ answers. The recordings were analysed a second time and were compared to the transcribed version to ensure that no information was left out from the original audio version. Also to facilitate easy and convenient access to data, all the interview transcripts were compiled in one Word document. The whole text containing teacher interview transcript was divided in seven sections corresponding to the seven teacher interviews. Similarly, the transcript of the students’ interviews was divided in four sections

corresponding to the four focus groups. The title of each section included the code of each focus group (A,B,C,D) or the initials of the pseudonym of the interviewed teacher, the class level and code of the school. Information in each transcribed interview was arranged following the order of questions as they appeared on the interview schedule. The prepared row data was then coded using manual nVivo Coding methods where the respondents' information was labeled using original wording.

Interview transcripts excerpt	Codes categories
<p><i>Researcher:</i> Have you assessed in this week? By assessment I mean tasks that you give to students and which you may or may not mark them. <i>PL:</i> yes I have..., <i>Researcher:</i> May I know what kind of assessment you gave? <i>PL:</i> written quizzes... and exercises <i>Researcher:</i> Ok, and what was the purpose of your assessment? <i>PL:</i> It was to see...if students... what students...after the lesson, what they gained from the lesson <i>Researcher:</i> What they gained <i>PL:</i> Yes <i>Researcher:</i> Ah... did you use rubrics? Rubrics I mean...can be a separate document that provides more details about the assessment... <i>PL:</i> I use rubrics when it is a test out of 100 points. But if it is out of 10 or 15 points I don't use it <i>Researcher:</i> That test out of 100, when is it given, is it after... <i>PL:</i> This is given at the end, after..., it's done once in a term <i>Researcher:</i> Ok so when you use it, do you share it with students? <i>PL:</i> it is given through instructions on the test paper, how they will do the test..., the number of questions etc <i>Researcher:</i> ok now about the methods you use to assess. Which ones do you normally use? <i>PL:</i> Ok you can use...we use ...question and answer: you ask and one student answers, you can... for example I've just used pair work, in pairs of two using one answer paper...yes... both written and oral are used <i>Researcher:</i> What do you mean by oral? <i>PL:</i> asking students in the classroom and they answer</p>	<p>(Written assessment) “written quizzes... and exercises” “what they gained” (Assessment of Learning) (Use of rubrics) “rubrics when it is a test” (Summative assessment) “given at the end” “given through instructions” (Use of Rubrics) (Oral assessment) “question and answer” “Pair work” (Grouping) “answer paper” “written” “oral” “Asking students”</p>

Figure 3.9: Sample of codes categorisation (Transcript extract from interview with Teacher PL)

The first cycle analysis of the recorded interview was done using content analysis methods which consists of extracting the original terms used by the respondents for interpretation i.e. manual nvivo coding (Saldana, 2009). The transcribed interview was coded by mining the respondents' key words to help identify the major categories that would enable full understanding of what the respondents were focusing on in their responses (See appendix 11). Figure 3.8 shows the sample of initial coding of the interview with Teacher PL. The major concepts and ideas were identified from the respondents' answers for each question. The

interview questions sought detailed data on when and why teachers assessed their students, the everyday methods that they used to assess, whether and how they provided feedback and other useful information that could enable me to understand every aspect of their classroom assessment practices. The codes were generated in such a way that they reflected the main ideas on these question topics.

After data was sorted and codes generated in the first cycle of the coding stage, the second cycle was carried out to identify the codes categories and major themes. *Pattern coding* methods were used to group together terms that referred to one common and meaningful construct (Saldana, 2009). As highlighted in Figure 3.9, codes pertaining to the same idea were highlighted in the same colour and a different name of the category was created to define the grouped terms in one concept (see Appendix 11). Codes were also grouped in sub-categories that helped to generate the main categories. These categories were also mapped to questions which in most cases were formulated to collect data on each topic (See Table 3.2). The topics were therefore the main units of analysis from which a detailed description of what teachers did during classroom assessment was developed. Overall, four main themes were described in this study namely the teachers purpose of assessment, the forms of assessment that teachers used in their classroom, their methods of providing feedback and the source of assessment tasks.

3.8 Ethical considerations

In this study, the research ethics was one of the major preoccupations of the researcher and necessary measures were taken to ensure that all ethical considerations required for this specific study were met. Research ethics is defined as “a matter of principled sensitivity to the rights of others” (Cavan, as cited in Cohen et al., 2007, p. 58). Although every research purpose is to uncover truth and presents facts, Cavan indicates that the dignity of those involved in the study should not be compromised by such quest for truth. Cavan’s definition highlights the researcher’s obligation to ensure that respondents are treated with respect and have to be given their rights even when such rights may be compromising. There is a double obligation facing the researcher therefore, one of conducting research in strict observation of ethical rules and another of not infringing to the rights of informants. Respect to the rights of those involved in

the research calls for refers to different obligations that set limit on what researchers can do and say to informants.

Ethical aspects of research also require conformity to rules of conduct defined by codes or a set of principles governing an institution (Israel and Hay, as cited in Robson, 2011). In the current study, I sought to meet the ethical requirements set by the learning institution affiliated with me i.e. the University of York. I also had to go through the process for obtaining research authorization from the ministry of education in Rwanda where data was collected. Before going for data collection, I obtained the ethical clearance from the University of York. This clearance was provided after all stages of data collection process were reviewed to ensure that the entire research activities were to be conducted in accordance with research regulations and guidelines in place at the University of York. In its article 11.2 on research integrity and ethics, the policy on research degrees at the University of York stipulates that the research student is responsible for securing the ethical approval before the start of research (University of YorK, n.d.).

In addition to the university ethical approval, I also had to obtain the permission from the office of research directorate in the ministry of education to certify that my research met the ministry research requirements before I could be allowed to conduct the research. The application process for the ministry authorization was forwarded to the Institutional Review Board (IRB) at The College of Medicine and Health Sciences at the University of Rwanda. The process involved the preparation and submission of an application package including copies of all instruments that were to be used in data collection and informed consents for all informants. An appointment was also arranged for me to attend a 30 minute presentation of my research project during which members of the IRB asked questions for clarifications and provided further recommendations. The permission was granted after my application was reassessed to ensure that the IRB recommendations were met. One of the key recommendations made by IRB members was to include a parents' consent form for every student participating in classroom observation and interview. As my study involved children under twenty years of age, I was required to have their parents sign consent forms. Students from schools where classrooms were observed were given the parents consent forms (See appendix 8) to bring to their parents for approval and signature. This meant that observation was possible only in day schools where students could go back home after school. In fact, all classrooms that were

observed in this study were from day school. All this was done to ensure that the data gathering process was conducted in accordance with the IRB recommendations and that research standards in relation with the respondents' rights to privacy, confidentiality and anonymity of identity were respected.

Diener and Crandall (1978) define privacy from three perspectives: (1) The information sensitivity which refers to how personal the given information is, (2) the setting being observed i.e. for example how private the venue is and (3) dissemination of information which consists of linking data to the identity of the informant. In this study, no activities that could infringe on the teachers or students' privacy were part of the research. All questions in the questionnaire and in interviews were centred on the usual teaching and learning practices and no private information were required. The research venue was usually classroom for students and mostly offices for teachers. Also, the information that was gathered was anonymous and there was no possibility of matching individual responses to the identity of the respondent. Furthermore, the purpose of the study and the whole process and procedures that were used during data collection and data analysis were all communicated to all participants so that anyone feeling that his or her privacy was threatened could be able to withdraw from the research.

The students signed assent forms (see Appendix 3) and all participating teachers signed consent forms as evidence that their participation was voluntary and that they fully understood what it took to be part of the research. According to Diener and Crandall (as cited in Cohen et al., 2007), informed consent is the procedure by which the would-be informants decide whether to participate or not to participate in a study after they are informed about the study and understand how their decisions might be influenced. In this study, all research activities were carried out keeping in mind my responsibility as a researcher to maintain the welfare of those involved in the research. This is in line with non-maleficence and human dignities values that ought to characterise a researcher (Cohen et al., 2007). The non-maleficence principle is based on the respect of the safety of the human subjects and avoiding doing any harm to them (Hartas, 2015). Both teachers and students were treated with honesty and dignity.

All participants were also given time to ask about the research and were informed about their right to withdraw from the investigation any time they wished to do so (Wilson and Fox,

2013). In addition to the written consent forms that provided detailed information about the study, I explained who I was, telling the students that I was a student too like them and letting teachers know that I was once a teacher like themselves. This helped to balance the ‘power relationship’ between researcher and participant which can sometimes affect data when it is not given enough consideration (Randal et al. 2013). In this context, the power relationship could be based on the differing level of education and the positions between the respondents (secondary school teachers) and me (lecturer at the university). For the respondent’s confidentiality, all data collection instruments were anonymised to remove any respondent private information such as names, telephone numbers or addresses. All information including data gathered by use of video tapes and audio recording were used for the sole purpose of this research and were in no circumstance shared by other parties. All the collected data were recorded and kept in password protected files only on my personal computer. Any collected information in form of video, audio or written data was destroyed after the analysis was completed approximately eighteen months after the completion of data collection process.

3.9 Summary of the chapter

Chapter three has presented the research methodology that has been used to gather and analyse data to answer the two main research questions. The chapter describes the size and the characteristics of research participants showing that 1258 male and female students and 66 teachers participated in the study. These were recruited from 41 schools located in five districts of the Southern province and three districts from Kigali city. The quantitative data was collected from teachers and students by use of the questionnaire and classroom observation schedules while the teachers’ interviews and students’ focus groups were used to gather qualitative data. The SPSS and thematic analysis methods were used to analyse the quantitative data and qualitative data respectively. After the pilot study which saw the number of the questions items reduced and some items wording changed to facilitate understanding, the main survey was carried out by the researcher in collaboration with teachers and or headteachers. The questionnaire return rate was satisfactorily high (92%) mainly because the researcher personally administered the questionnaire at most schools and travelled to each school to collect the completed copies in person. In the next chapter, the results obtained from the collected data are presented and analysed.

Chapter 4: The teachers' assessment practices and the students' perceptions

4.0 Introduction

This chapter presents the results of the analysis of the qualitative and quantitative data gathered for research question one: *'What are the classroom assessment practices of teachers of English and how do the students perceive these assessment practices?'* Data for this question was collected from both the teachers and the students and was intended to provide answers to five sub-research questions generated from research question one: (1) *What forms of assessment do the teachers use in their classroom?* (2) *For what purpose and at what time do the teachers use assessment forms in the classroom?* (3) *What is the content focus and source of assessment tasks used in the classroom?* (4) *What methods of providing feedback do the teachers use?* and (5) *What are the students' perceptions of the teachers' classroom assessment practices?*

The results presented in this chapter were organised by sub-research question and data collected through questionnaire surveys, interviews, focus groups and classroom observations for each sub-question are all reported together. In the same order as data was collected, the presentation of results on each sub-question begins with data from the questionnaire survey and then supported by data from classroom observation, interviews and focus groups. This order was maintained to facilitate the understanding of quantitative data and to help in the comparison and cross verification of information. The teachers' self-report from the questionnaire survey was compared to the researcher's own observations from the classroom and to the responses from teachers' interviews and students' focus groups. The patterns and relationship between these data were explored in order to provide answers to the relevant research question (Cohen et al., 2007). For sub-research questions where some data about the teachers' assessment practices was collected from teachers and students, data from both respondents was presented together for data triangulation purposes. Both teachers and students were asked to indicate how frequently different assessment forms were used in the classroom and the obtained sets of data from the questionnaire, observations, interviews and focus groups

were all compared for consistency. Data on the teachers' classroom assessment practices was presented in this chapter as a way of leading in to the detailed analysis of the students' self-efficacy in chapter five.

Tables, charts, figures and interview comments are used throughout the chapter to present the results on both research question one and research question two. Descriptive statistics was used to describe data on teachers' assessment practices mainly the forms of assessment, methods of providing feedback and the students' perceptions. Overall, the quantitative data in this chapter are presented by use of descriptive statistics mainly frequencies, modes and mean averages to show the extent to which specific practices were used for classroom assessment. Numerical data is accompanied by interview transcript excerpts and or comments put forward by the respondents during interviews and highlighting key points pertaining to each sub-research question.

4.1 Forms of assessment used in the classroom

This section presents results for sub-research question 1.1: *What forms of assessment do the teachers use in their classroom?* The results demonstrate how frequently different forms of assessment were used in the classroom. It might be useful to restate that quantitative data that was collected to answer research question one derived from a questionnaire survey which involved 66 teachers and 1258 students from 41 schools. Relevant qualitative data was also gathered from seven classroom observations, seven teacher interviews and four students' focus groups. More details about the participants are provided in Section 3.5. Information on the forms of assessment that teachers used in their classrooms was gathered mainly by means of self-report because there were few sources available at the surveyed schools that could provide information on the teachers' common approach to language assessment.

In addition, the scarcity of English textbooks was cited by the respondents as one of their big challenges (see Section 4.2.4) which compelled many teachers to find sources of their assessments on their own. Some textbooks distributed by MINEDUC through REB were available at some schools. These were mainly grammar-based but also contained some

assessment activities pertaining to the four English language skills. At the moment of data collection, no classroom assessment guides were available to the surveyed teachers of F2 and F3 of the lower secondary schools. Gathering information from teachers and students seemed to be the appropriate way of knowing what happened in the classroom in relation to assessment. This section presents the results obtained from both the teachers and the students on the different forms of assessment that were used in the classroom specifically between *paper and pencil based* and *performance based* forms. As is explained above in the introduction of the chapter, the results from the questionnaire survey are presented first and supported with the results of the analysis of qualitative data from classroom observation, teachers' interviews and the students' focus groups.

4.1.1 Paper and pencil versus performance assessment

In the current study, two terms: '*performance assessment*' and '*Paper and pencil*' (see Section 2.2.4) were used to classify the forms of classroom assessment that were recorded during classroom observation or self-reported by teachers. Rhodes et al. (1996) explain that '*paper and pencil*' forms of assessment refer to more controlled methods (e.g multiple choice questions) whereas '*performance assessment*' is defined as assessment using freer methods such as oral debates and interviews. In the questionnaire survey, teachers were asked to indicate how frequently they used different assessment forms in their classrooms within one school term. A list of fifteen forms of assessment (see Table 4.1) was provided showing items related to both '*paper-and-pencil*' (highlighted in italics in table 4.1) and '*performance*' based assessment forms. For example, the 'multiple choice' items were classified as relating to '*paper and pencil*' based assessment while other assessment forms such as 'essay writing' and 'oral presentation' are relevant to '*performance*' assessment forms. Table 4.1 indicates the teachers' self-reported answers on the forms that they frequently used in their everyday classroom assessment. The table displays the modes and mean averages (in a descending order) of frequency per term that each form of assessment was used on a range from 1 (never used) to 4 (usually used).

Table 4.1: Forms of assessment used by teachers in the classroom (Teacher self report)

	N	Never (Never used)	Rarely (once a week or once every two weeks)	Sometimes (two to three times a week)	Usually (more than three times a week)	Mode	Mean	Std. deviation	
		1	2	3	4				
		%	%	%	%				
1	<i>Sentence completion</i>	48	0	6.3	43.8	50	4	3.4	0.6
2	<i>Short-answer items</i>	45	2.2	8.9	40.0	48.9	4	3.2	0.7
3	Reading aloud (for reading fluency)	46	0	13	39.1	47.8	4	3.2	0.7
4	<i>Multiple choice</i>	56	1.8	26.8	30.4	41.1	4	3.1	0.8
5	Book reading (for reading comprehension)	42	7.1	19	31	42.9	4	3.1	0.7
6	Discussion/debate	45	6.7	17.8	53.3	22.2	3	2.9	0.8
7	Oral presentation	44	4.5	29.5	34.1	31.8	3	2.9	0.9
8	<i>Matching/ true-false</i>	42	7.1	26.2	38.1	28.6	3	2.8	0.9
9	Paragraph writing	45	8.9	28.9	40	22.2	3	2.7	0.9
10	Essay writing	48	10.4	39.6	31.3	18.8	2	2.5	0.9
11	Summary writing	48	8.3	43.8	29.2	18.8	2	2.5	0.8
12	Role play	35	20	37.1	20	22.9	2	2.4	1
13	Student Interviews	32	31.3	25	25	18.8	1	2.3	1.1
14	Poem writing	32	56.3	25	15.6	3.1	1	1.6	0.8
15	Listening to audio	34	70.6	17.6	5.9	5.9	1	1.4	0.8

Teachers selected their most used assessment forms and indicated the frequency of their use by choosing an answer on a four Likert scale from 1= *Never* to 4= *Usually*. In order to avoid different interpretations that the respondents may have of the scale, supplementary description was added to clarify the meaning of each scale point provided on the questionnaire where ‘*usually*’ was defined as ‘*more than three times a week*’, *Sometimes* as ‘*two to three times a week*’ and *Rarely* defined as ‘*once a week or once every two weeks*’. Choosing *Never* meant that teachers “*never used it*”. As illustrated in Table 4.1, up to 43.8 % of teachers who responded to the questionnaire indicated that they used ‘*sentence completion*’ at least two to

three times a week while half of all teachers reported using ‘*sentence completion*’ forms more than three times a week. In addition, The lowest standard deviation of 0.6 and the highest mean ($\bar{x}= 3.4$) also indicate that most teachers usually used ‘sentence completion’ forms, which is part of ‘*paper and pencil*’ based assessment, for their classroom assessment. Table 4.1 also shows that a high percentage of respondents (70 %) reported that they never used listening based forms of assessment. This was also the case for ‘poem writing’ and ‘student interview’ which both have low means of $\bar{x}= 1.6$ and $\bar{x}=2.3$ respectively, thus suggesting that the majority of teachers rarely used these two forms of assessment in their classrooms.

Overall, the results in Table 4.1 suggest that the majority of the teachers used *paper and pencil* (shown in Table 4.1 in italics) related forms. In fact, of the top five frequently used assessment forms, three of them (*multiple choice, sentence completion and short answers*) belong to *paper and pencil* based assessment. Figure 4.1 indicates the comparison of the mean average between the *paper and pencil* based forms (item 1, 2, 4, 8) and *performance* based assessment forms for each language skills i.e. for writing (item 9, 10, 11, 14), for speaking (item 6, 7, 12, 13), for reading (item 3, 5) and for listening (item 15).

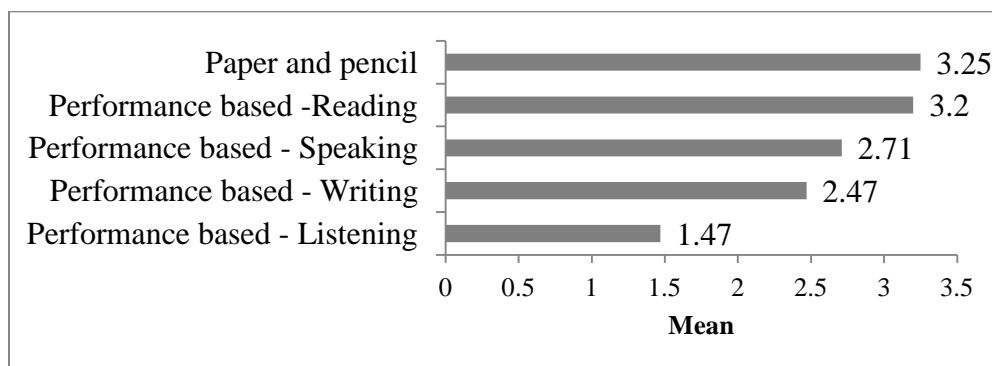


Figure 4.1: Frequency of use of Paper and pencil vs performance based assessment forms

As illustrated in Figure 4.1, the teachers’ preference for ‘paper and pencil’ based assessment is shown by the highest aggregate mean average ($\bar{x}=3.25$) of frequency for the 4 first items relating to ‘paper and pencil’ assessment. Although it was anticipated that ‘paper and pencil’ based assessment would be reported as more frequently used than ‘performance’ assessment, it appears that the frequency of use of ‘performance’ based assessment, as reported by teachers, was higher than expected ($\bar{x}=3.2$ and $\bar{x}=2.7$ for reading and speaking respectively). In addition,

Table 4.1 indicates that more than 75 % of teachers said that they used ‘debates’ or ‘discussions’ at least two or three times a week which seems to be high given that studies have consistently indicated a much less use of performance assessment forms in many similar EFL classroom contexts (Tsagari and Pavlou, 2009; Zhang and Burry-Stock, 2003; Mertler, 1998). For purposes of data verification and data triangulation, more data on the forms of assessment used in the classroom was collected from students. They were asked to indicate the extent to which different forms of assessment were used by their teachers in their classrooms (see Appendix 1A for students’ questionnaire) and the results were compared to the teachers’ report.

Data from students on the assessment forms that teachers used in the classroom assessment partially agreed with data from teachers’ self-report. Similar to the results from the teachers’ self report, the results from the student questionnaire indicated that the forms of assessment related to ‘*paper and pencil*’ such as ‘gap filling’ and ‘multiple choice’ were the most used in the classroom. As shown in Table 4.2, the students’ responses were consistent with the teachers’ self-report, with the highest mode (4) and mean ($\bar{x}=3.4$) for ‘paper and pencil’ based assessment forms. In addition, the calculated measurement of data dispersion for both teachers’ and students’ responses on ‘*paper and pencil*’ based forms of assessment (in Table 4.2 highlighted in italics) gives the lowest standard deviation point ($SD=.6$) compared to the responses on ‘performance’ based forms. This highlights the existence of some consensus among respondents on this question. The results in table 4.1 and table 4.2 also demonstrate consistency in how listening assessment forms remained the least used (Mode =1, $\bar{x}= 1.5$). On the other hand, however, the students’ responses also indicated that ‘*performance*’ based assessment forms were less frequently used than was reported by teachers. For example, unlike the teachers’ self report in Table 4.1 which shows that only 21% of teachers affirmed having never used ‘role plays’, Table 4.2 indicates that more than a half of the students reported having never had assessment based on ‘role plays’.

Table 4.2: Students reported forms of assessment used by teachers in the classroom

	N	Never	Rarely	Sometimes	Usually	Mode	Mean	Std. Deviation
		1	2	3	4			
		%	%	%	%			
Written based assessment (e.g: fill in questions, underlining right or wrong answer, multiple choice questions, etc)	1246	1	7.8	38.2	53	4	3.4	0.6
Reading based assessment (e.g. reading comprehension test)	1252	10.2	13.3	28.6	47.7	4	3.1	1
Writing based assessment. (e.g letter writing, summary or paragraph writing)	1245	15	14.2	38.2	32.6	3	2.8	1
Speaking based test. (e.g. Oral presentations, debates or discussions)	1253	34.6	18.7	30.5	16.2	1	2.2	1.1
Role play based (e.g. Performing a written dialogue)	1239	52.7	16.5	22	8.9	1	1.8	1
Listening based assessment. (e.g. Listening to audiotapes or to the radio)	1242	66.4	15.5	13.9	4.2	1	1.5	0.8

Also, while the results in Table 4.1 suggests that more than a half of teachers reported using debates or discussion at least two times a week, Table 4.2 indicates that only 30.5% of the students agreed that debates and or discussions were used at least two or three times a week. This discrepancy in the students' and teachers' answers raises the question of the validity of data from self report where teachers might have wanted to show that they were aware of and used different forms of assessment. Nonetheless, the prevalent use of '*paper and pencil*' based forms was emphasised by the consistency of the findings from both the teachers and the students who indicated that performance assessment forms were less used than paper and pencil (see Table 4.1, Figure 4.1 and Table 4.2)

Similarly, the results from the qualitative data particularly from interviews with teachers and students focus groups seemed to echo findings from the teachers and students questionnaire survey regarding the forms of assessment and the frequency of their use in the classroom. The majority of interviewed teachers confirmed that their assessment was focused on controlled

assessment (paper and pencil) and their detailed explanations on how frequently they used assessment such as oral presentation and essay writing inferred that performance assessment was much less used than was reported in the questionnaire. This could indicate that the questionnaire data was not highly reliable as teachers seemed to have overstated their use of performance assessment in the classroom. Commenting on whether and how often they used specific tasks such as debates and individual oral presentation, most teachers explained that such tasks were rare and were only planned at specific times of the school term. For example, Teacher MR explained that it required her to plan extensively for oral assessment. Answering to my question on the form of assessment that she used more often between written assessment such as written tests, and performance assessment such as oral presentations or debates, Teacher MR stated:

It's usually written assessment [...] oral is given for example after one month of preparation after students finish reading books...it's time consuming... A week may end without oral assessment. Written assessment is the one mostly used. It may be used twice or three times a week or so. Oral ...I prepare it for example I can decide that I will have two oral assessments in one term... But they are very few.
(Teacher MR)

Responding to the questions of the research on the type of tasks and methods they used, the common answer was that teachers wrote questions on the chalkboard and students wrote their answers on a sheet of paper.

I usually use individual work...I write the questions on the blackboard and each student writes answers on his or her answer sheet[...]it also depends on the lesson, but the written assessment consisting of writing answers on a paper (paper and pencil) is the most frequent. Others such as role play are very rare. (Teacher PL)

Teachers also explained that 'paper and pencil' assessment that they used often consisted of *short answers* or *sentence completion* among other forms. Different factors such as time constraint were put forward which, according to the interviewed teachers, were behind the challenges that hindered the use of performance based assessment. The teachers indicated that preparing assessment such as oral based tasks was time consuming and required hard preparation both for teachers and students, which seemed to explain why they were not so frequently used. Most teachers said that they sent their students to the libraries to read books and orally present the summary in the classroom. This reportedly required more time compared

to written assessment that teachers were able to design even in the middle of a lesson (Teacher JS).

What is most frequently used is written [...] because oral you often realise that students say 'we have not prepared...'... and you'll notice that they have no ideas ... oral ...it is not used often because it requires you to give them work to bring home, for them to take enough time ... but written they do them every day (Teacher JS)

In addition to the time constraint that was commonly cited, teachers teaching in F3 also said that they did not focus on performance-based assessment because it was not assessed in the Lower Level National Examination taken every year by students in F3.

VN:...again why I.I can even say why most of the time I use written... eh written ...Because even national examination is written-based... The students do not express themselves orally.

Researcher: so you want to train them on how to pass the national test

VN: yeah yes I think that is what we do.

Writing assessment focused on summary writing was repeatedly mentioned as one of the writing tasks that teachers often used in the classroom. Even as it appeared that paper and pencil based assessment was the most frequent, all teachers said that they also gave reading and speaking assessment tasks. However, it appeared that some teachers did not have full understanding of the boundary between assessment of reading and assessment of speaking. For example, for Teacher VN, reading a written passage aloud or orally answering the teacher's reading comprehension questions meant that students were being assessed on reading and speaking skills at the same time. He considered oral questioning as his most frequently used oral assessment method that he used to check his students' speaking skills.

Researcher: So in general [...] what forms of assessment do you often use in your classroom?

T: What is frequently used is written assessment and oral assessment ...I mean it's often written assessment rather ...I'm sorry.

Researcher: [...] if you can remember, how many times have you used those written assessment in this week?

T: Like three times

Researcher: How about oral assessment?

T: [...] When I'm teaching I sometime ask students questions ...I think that is a kind of assessing...

Teacher VN comments above about oral assessment made me question the teachers' responses in the questionnaire on how frequently they used *performance* assessment. Teacher VN's comment equating 'teacher oral questioning' to 'performance assessment' could imply that some teachers might not be aware of what constitutes real assessment of each language skill. Teacher PL explicitly reported that he considered classroom oral questioning as his principal form of performance assessment. This could have affected some teachers' answers in the questionnaire resulting in the inflated frequencies of use of *performance* assessment at different points of time during instruction. As shown in Table 4.6 of Section 4.2.2, most teachers indicated that they used '*performance*' based assessment more frequently than '*paper and pencil*' both during and at the end of instruction. This is why the use of mixed methods was proved important and useful for the current study as it allowed for the comparison of different sets of data. Many teachers' responses during interviews helped to understand further the answers provided in the questionnaire.

In addition, data from classroom observation was used to cross check the validity of the teachers' answers in the questionnaire survey and from interviews concerning their most frequently used forms of classroom assessment. The results from observations (see Table 4.4) indicated that no real *performance* based assessment was actually used in any of the observed classrooms. This differed from data from the questionnaire survey which pointed to the teachers' frequent use of *performance* assessment. The analysis of data from the classroom also suggested that teachers' oral questioning was frequently used in the classroom. The results demonstrated in Table 4.3 show the number of occurrences of the teachers' oral questioning in the seven observed classrooms. Observation in each classroom lasted for an average of 40 minutes where a video camera was used to record the interaction between teachers and the students in the classroom. These results seem to suggest that the majority of the observed teachers spent considerable time using oral questions therefore not allowing enough opportunities for the students to engage in performance assessment such as debates and oral presentation for the speaking skill, or paragraph and essay writing for the writing skill.

Table 4.3: Teacher-student classroom interaction

		N*	Frequency
1	Learner answers a question	7	109
2	Teachers asks a display question	7	102
3	Teacher praises the student's answer	5	13
4	Teacher asks for more answers	4	12
5	Teacher explains a grammatical point	5	9
6	Teacher answers the student's question	3	7
7	Learners asks a question	3	7
8	Period of silence (no answer)	2	5
9	Teacher asks a referential question	2	4
10	Teacher comments on the student's response to provide further clarification	3	4
11	Teacher gazes to ask for clarification/ prompt student to answer loudly	2	3
12	Learner talks to another learner	3	3
13	Teacher explains a point relating to the content of the lesson/assessment	1	3
14	Teacher explains the meaning of a vocabulary	2	2
15	Teacher gives instruction/directions on assessment	1	1
16	Teacher gives assessment results	1	1
17	Teacher explains a functional point	0	0
18	Teacher criticizes the student's answer	0	0

N represents the number of classrooms in which a given incidence of interaction was observed

It was also observed that some teachers' oral questioning was accompanied by writing where teachers spent considerable time writing lesson content on the chalkboard for students to copy in their notebooks. Taking much time copying notes also meant that students had limited time to practice the learned language. As can be seen in Table 4.3, the number of occurrences of the students' responses was higher than the number of occurrences of teachers' oral questioning because teachers sometimes asked for more answers for one question. This often happened when teachers were not entirely satisfied with the students' answers.

The observed common use of oral questioning by teachers in the context of the current study was expected as several research studies have indicated that the majority of teachers in many educational settings use oral questions in their classrooms (Thornbury, 1996; White and

Lightbown, 1984). Also, having been a learner and a teacher of English in Rwanda and thus having witnessed the extensive use of a more or less teacher-centred teaching approach at the lower level Rwandan secondary school, it was not surprising to me that teachers used more display questions than inferential questions in their classrooms. In general, teachers asked closed questions which required short answers and students were often not asked to elaborate further as illustrated in the following interactions between teacher VN and his students.

- VN: Ok today we are going to start a new lesson; it is called ‘summary writing’ [...] what is a summary?
S1: Few words
VN: Ok. Few words (T writes ‘few words’ on the board) it means summary is to say in few words. What is ‘few words’? You are saying few words of what?
S2: a story
VN: Ok a few words of a story (T writes ‘story’ on the board)
S (shout outs): Text, Novel, poem...
VN: (Teacher writes the students’ answers) A few words of a Text, Novel, poem, theater, drama, newspaper...it means that you can do a summary of what is written or what is spoken [...]
S3 (asks in native language if news headlines can be considered as summary)
VN: Say it in English (another student laughs).
S3: Headlines... (Student hesitates).
VN (Interrupts): Headlines. They are summary. You are going to ask whether headlines are summary. Yes they are summary. Headlines of news are summary [...] I can say we have two types of...
S (chorus): Summary
VN: There is general and...
S4: Short
VN: And short summary [...] what’s the difference between general and short summary?
S5: General make many paragraph
VN: Ok when you are writing a general summary you make many paragraphs; Very good. For short stories what about short stories?
S6: Only one paragraph
VN: Ok for short story you make only one ...
S7: Paragraph

It shows from this extract that the students’ answers to the teachers’ oral questions were short and sometimes given simply to complete the teacher’s sentence. The teachers’ use of incomplete sentences was noted in all the observed classrooms. The teachers’ questions were formulated in the form of incomplete statements where students were required to provide a missing word (see S4 and S6 above). However, unlike Teacher VN, some teachers like Teacher

FR also encouraged their students to provide full answers emphasizing on the use of complete sentences:

FR: What does 'paraphrase' mean [...]? Yes (teacher points at a student) try
S1: is...
FR (interrupts): Don't start by the verb. Start with the subject if it is there, verb and then add an object
S1: paraphrases
FR (interrupts): Paraphrase is ...
S1: is the word
FR (interrupts): No repeat: paraphrase is...
S1: Paraphrase is the word which has ...which has the same meaning and ... is different...to writing
FR: Good. Clap for her [...] who can add something?
S2: paraphrase...
FR: A paraphrase...
S2: A paraphrase is a change of forms, sentence but not meaning

In some cases, teachers encouraged the students to ask questions but students often appeared reluctant to raise their hands for questions. Some students also often chose to ask questions in their native language and appeared to be uncomfortable when the teacher asked them to speak in English. The students' tendency to speak in their first language was not surprising because it was the language that they used in their everyday communications when they are not at school. It was also noted that some teachers used the students' first language to provide clarifications in the classroom in spite of the fact that classroom was perceived to be the only place for many students to speak English in a purely monolingual society where more than 99% of the population speak Kinyarwanda (MINECOFIN, 2005) which is the only national language. However, the students' limited proficiency in English seemed to obstruct their participation in the teacher-student interaction. Although it seemed that teachers used oral questions to open up questions to more students and to check their understanding, the fact that many of their questions were display questions only allowed for a very limited student participation and offered little opportunity to practice the language.

In classroom assessment, oral questions are usually used to check for learners' understanding (Ma, 2008; Ozcan, 2010), to get the students attention (Chaudron, 1988) and can be used to motivate students and encourage them to participate in the classroom learning process and help them use and apply their newly learned ideas (Barnes,1990; Khalk et al., 2014). It could be argued, however, that although the teachers' use of oral questions seemed to

be aimed at encouraging students' participation and gathering information about their understanding, it appeared that the students' answers were usually short and could hardly reveal the level of learners' understanding.

Some teachers did not provide constructive feedback when wrong answers were given. Instead, more students were prompted to answer until a correct answer was obtained. While these findings are revealing of the teachers' assessment practices, they may not be generalized to the whole ordinary level of secondary school in Rwanda. The current data was only obtained from day schools due to challenges met when attempts were made to observe classrooms in boarding schools (see Section 3.3.4). The researcher was asked to obtain the parents' signed consent before the students could be allowed to be observed or before they could participate in focus groups. However, it was impossible to reach every student's parents because students in boarding schools came from all parts of the country. The second option was to ask the students in the surveyed boarding schools to travel home and get the parents' consent forms signed but such permission could not be granted. For these reasons therefore, the results from classroom observation should be generalised with caution as these described limitations may have affected the validity of the collected data.

Table 4.4: Classroom assessment events

<div style="display: flex; justify-content: space-between; align-items: flex-start;"> <div style="width: 20%;"> <p>g/p Group/pair work,</p> <p>x Teacher oral questioning (Teachers asks oral questions and students answer)</p> <p> Paper and pencil assessment task ,</p> <p> Performance assessment task,</p> </div> <div style="width: 20%;"> <p>i Individual assignment ,</p> <p> Teacher explains / writes on the board ,</p> <p> Correction,</p> </div> <div style="width: 20%;"> <p>m Teacher marked work,</p> <p>m/f Teacher monitoring & feedback</p> <p>h Students write answers on the board,</p> </div> <div style="width: 20%;"> <p>u unmarked work</p> <p>n Teacher uses native language to explain</p> </div> </div>												
Tea- cher	Time line (minutes)											
	1-5 th	6-10 th	11-15 th	16-20 th	21-25 th	26-30 th	31-35 th	36-40 th	41-45 th	46-50 th	51-55 th	56-60 th
PL	x x x	x x x	x x x	i u	i m/f	i x b	x b	x b	x p m	p m	p m	x b
FR	x x	x x	x x	x x	n x n	x x	x x n	n	x x	x x	x n n	
MR	x	p x	b	g g	m/f g	m/f g	b x b x b					
VN	x	x	x			x		x				
JS	x x	x	x x	x x	x x	x						
SF*		x	x	x	x	x	x	x	x	x	x	x x
SL	x	x	x	x	x	x	x	x b	x b	x	x x	g m/f b b

*Teacher SF’ session was focused on correcting a test that was given prior to observation

Table 4.4 indicates that paper and pencil based assessment was done in four out of seven observed lessons. It also shows that oral questioning was widely used by teachers both during explanation time and during assessment tasks and correction. Six lessons were focused on grammar apart from that of Teacher VN whose lesson was focused on summary writing. Teachers SF lesson was about correction of a test he had given prior to the classroom observation. As was also emphasized by some teachers during interviews, group work seemed to be more frequent than individual work. In addition, it was observed that the teachers usually monitored and provided support to students only when the students were doing their tasks in groups.

The teacher-centred approach which seemed to dominate in all the 7 observed classrooms could be viewed as an explaining factor to the lack of the students' full involvement and active participation in language practice activities. Lessons in all observed classrooms were dominated by teacher talk and all teachers seemed to follow the same teaching pattern. First, the majority of teachers introduced the lesson by telling the students what the day's topic was going to be, wrote the title on the chalkboard and started asking display questions as they explained. For example, Teacher FR started with examples and asked leading questions that required students to find out the topic of the lesson. Teacher JS began with revision, asking students to recall what had been taught previously. Secondly, the teachers explained the topic, using many oral questions to check understanding. Most learners sat and listened to the teacher's talk and answered to the teacher's question when singled out. Most of the students also copied what teachers wrote on the chalkboard as he or she explained. The teacher-student oral questioning occurred at anytime during lessons. Teachers also often addressed their questions to the whole class for volunteers to raise their hands and give responses. Some teachers also singled out individual students and asked them to answer. In addition, most teachers integrated some students' answers in their notes by asking students to write them on the chalkboard for other students to copy in their notebooks.

Data from classroom observation was very revealing as it yielded real data from real classroom environments. Information presented in Table 4.3 and Table 4.4 provided some useful insights into how teachers assessed their students and allowed for more questioning of data from the questionnaire survey and interviews. However, it is important to acknowledge that the short observation time and the limited number and range of observed lessons may not be enough to get a full picture of how classroom assessment was carried out. Longer observation time in a wide range of classrooms could have provided more varied data and could be more appropriate for generalisation.

In addition to the results from the teacher questionnaire survey, teacher interviews and classroom observation, the analysis of data from the students' focus group discussions confirm the prevalent use of 'paper and pencil' based assessment forms. The participants indicated that they were more familiar with reading and grammar focused forms of assessment than assessment focused on practice such as debates and oral presentations. As echoed in Group A

and B extracts below (See the full extract in Appendix 10), grammar and vocabulary based assessment was reportedly frequently used in the classroom while some performance based activities specifically debates were said to take place on the weekends.

GROUP A

S1: We focus on texts and summaries. They give us a text then you read it and summarise it in your own words.

S2: We also do vocabularies

GROUP B

Researcher: So you all seem to agree that the content of the assessment activities is always related to the learning content.

S (Chorus): yes/right

Researcher: Right...so my question then is: What is the focus of your learning content?

S1: Verbs

S3: Tenses

Researcher: Ok ...verbs and tenses... anything else?

S4: Nouns

Researcher: That is grammar ... right? How about essay writing? Or writing stories, letters...how often do you do that?

S (chorus): No, we don't do it.

Researcher: You don't do it?

S1: No

Researcher: How about speaking? Like learning how to do debates etc...?

S2: We do it. Every Friday a debate is always scheduled

Researcher: Ok...and that is every week?

S2: It is once a month. The last Friday of every month between students of different class levels

Researcher: I see...how about in your own classroom? Does it happen that the teacher can notify you that you will learn about speaking? Say ...that you will learn about how to organise an oral presentation in the classroom?

S5: No... that is very rare...we don't do it.

Researcher: Ok...so it's about grammar and assessment is usually focused on grammar as well.

S (chorus): yes

It was also observed during interviews in focus groups that some students seemed to confuse assessment on writing and written assessment. For example, some students' comments seemed to suggest that any written task that they did in the classroom was part of writing assessment. Although I often provided good examples to the students at the beginning of each interview of how writing, reading, speaking and listening assessment can differ based on learning objectives, some students still tended to say for example that a grammar based quiz was also

about writing, or a short text reading was also about speaking because sounds were produced while reading aloud. In order to provide more clarification to the respondents and thus to increase the reliability of their information, follow up questions were asked and specific examples were used in the case of confusion.

4.1.2 The use of rubrics

Data on the use of rubrics were collected to help understand the contexts in which the assessment forms were used. Although rubrics are not used in the same way as assessment forms, their use serves a complementary role to assessment and benefits both teachers and students. Data on rubrics was related to research question one and was sought to explore the extent to which their use enhanced the transparency of assessment. Also, as explained earlier in Section 2.2.4, the frequent use of performance assessment would likely require the use of rubrics (McTighe and Ferrara, 1998). Information on rubrics was therefore gathered as additional evidence indicating the possible use or non use of assessment forms. In the current study, the results show that 86% of the interviewed teachers indicated that they had never used rubrics during assessment. In most cases, assessment that was given during or immediately at the end of the lesson was often unannounced and not preplanned and no rubrics were used to provide details about what students were expected to achieve on the assessment. This may have arguably led to the observed students' misunderstanding of the nature and purpose of some assessment tasks.

The use of rubrics is generally recommended for providing details about the criteria related to learning that students are required to demonstrate and clarifying the quality of performance that students should aim to attain (Brookhart, 2013). These rubrics can be shown to students before assessment tasks so that they can know what the assessment goals are (Suskie, 2009). As was explained to the respondents, the term rubrics in this study was used to refer to a set of criteria which can be used to inform learners about the quality that their work should have and which guide teachers for during the process of measuring the quality levels of the students' performance (Brookhart, 2013). For the assessment of essay writing for example, the criteria in the rubric may be to provide good examples, choosing the right audience, using rhetoric devices to name a few (Doghonadze, 2017). The results of the current study showed

that teachers were not conversant with the use of rubrics. Some teachers even expressed their disapproval of the use of rubrics. In this regard, Teacher SL argued that in ‘learner-centred approach’, students should be able to figure out on their own what they have to do during assessment tasks on their own.

Because our current method is learner-centred, a student being the centre of knowledge, we don’t tell the student that you will do this way immediately. If I have explained clearly the question, you write the question, and you see if the student discovers it. You write the questions and if they say that they don’t understand this or that, in that case you explain and say ‘you will do this way’, but if you tell them clearly that ‘if you do this you will be marked this way..’ No that can be problematic. (Teacher SL)

Although teachers claimed that they used oral instructions to ensure that students understood the questions, it was found that most teachers never used rubrics and some had no knowledge of how rubrics work. Teacher MR explained that she used written rubrics to help her remember how to mark her students especially when assessment was about reading and oral assessment.

I prepare a marking scheme first..., it helps me to know if this student forgot to do this..., this amount marks will be taken off..., or if they ask whether they would get marks... if they write one thing or another then I know what to tell them. I write it in question papers on the quiz preparation sheet. If it is a quiz for example...or about reading or ... if I decide to give oral assessment; In that case, I write it on a separate paper and explain that a student who will do this or who will read with ‘eyes contact’ ...I will add some marks... If a student raises their voices, etc. The student who will pronounce this way ... all this get written down on a paper. (Teacher MR)

Teacher MR’s views were echoed in Teacher SL’s comments who pointed out that ‘rubrics’ were not communicated to her students before assessment. However, she explained that some clarifications regarding the marking procedures were sometimes provided.

I don’t tell them before. What I often tell them before the assessment is for example that if they do not get 50% score, some marks will be deducted from their previous score. If for example a quiz is out of 10, I tell them that if they don’t score 5, then 2 marks will be deducted from their score in the previous assessment. They therefore work accordingly. But I don’t tell them that if they lose marks ... if they don’t get the tense correct..., or if they forget an object... No. I don’t tell them that. I usually plan for it but I don’t tell them. (Teacher SL)

Instructions written on blackboards or on question papers and brief verbal instructions provided during classroom tasks were the two main methods through which students were informed about what they had to do during assessment.

Teachers' perceptible lack of interest in rubrics and their apparent lack of knowledge with regard to the use of rubrics may be related to their preference of *paper and pencil* assessment. It has been argued that rubrics are more relevant and highly essential for performance assessment where students' need to be clearly informed about the achievement criteria upon which their performance is assessed (Andrade, 2000; McTighe and Ferrara, 1998). Overall, the teachers' self-report indicated that they used *paper-based* assessment more frequently than *performance-based* assessment in the classroom. This prevalence of paper and pencil based forms was also noted during classroom observation where, as was also reported by some teachers during interview, assessment tasks were often unannounced and usually written on chalkboards. In many cases, these were collected by teachers sometimes to be assigned marks that would count for the students' official performance report at the end of the term. Written tests were reported to be highly weighted and used mainly at the end of an important component of the lesson or at the end of the school term. On the other hand, the majority of the teachers agreed that performance-based assessment was less frequently used in their every day classroom assessment except for the unmarked teacher's oral questioning. Teachers also reported to use well planned and marked performance assessment such as debates and oral presentations at designated time in a term. Although this sounded a reasonably good practice, the frequency of those performance assessment tasks was very low compared to the common paper and pencil assessment. For example, Teachers MR, JS and SL demonstrated that performance assessment could sometimes occur once in a term due to the time constraints and other challenges associated with their preparation.

4.2 Other assessment practices

In this section, more results on the teachers' classroom assessment practices are presented. These are results from data on the usual purpose of teachers' assessment, the time at which the teachers usually assess their students, the content focus of their assessment and the sources of

their assessment tasks. This was done in order to explore further other aspects that characterise the teachers' classroom assessment and to understand the context in which different forms of assessment were used. The analysed data was related to sub-research questions 1.2 and 1.3 and was collected mainly from the teachers during the questionnaire survey. Some data was also gathered during interviews with teachers. Information on other classroom assessment routines was sought to help understand the realities surrounding the teachers' every day classroom assessment and provide explanations on the use of some assessment forms. This information was also used to contextualise the students' perceptions and self-efficacy beliefs. It was used for the interpretation of the students' answers on other research questions particularly in relation to their perceptions of their teachers' classroom assessment practices as a whole. The following sections reports on the results from the teachers' and students' questionnaire survey supported with data from teacher interviews and students' focus groups with additional comments from classroom observations.

Data on other assessment practices was analysed using SPSS descriptive statistics. These consisted mainly of the calculation of percentages and mean averages in order to identify the most preferred purpose of assessment for which teachers assessed their students. Descriptive analysis also helped to identify the specific times of instruction at which teachers carried out their classroom assessment. Percentages and mean averages were also used to know the specific content on which the teachers' assessment was usually focused as well as the common sources of assessment tasks used in the classroom. These results are presented in tables 4.5, 4.6, 4.7 and Table 4.8 as shown in the following sub-sections. The statistical analysis was also supported with the findings from interviews and notes from classroom observations.

4.2.1 Purpose of teachers' assessment

Part four of the teachers' questionnaire was reserved for gathering data on the purpose of classroom assessment which, as some argue, is one of the factors influencing teachers' choice of tools to be used for classroom assessment (Chapelle and Douglas, 2006). Information collected through this question item was thus intended to help understand why teachers assessed in the way they did in their classrooms, and to explore the possibility that the teachers' purposes of assessment influenced their choices of assessment forms.

Table 4.5: Teachers' purpose of assessment

Purpose of assessment	N	Never	Rarely	Sometime	Usually	Mode	Mean	Std. deviation
		1	2	3	4			
		%	%	%	%			
1. Provide feedback to students in order to improve their learning process.	63	0	0	11	89	4	3.8	0.3
2. Measure extent of learning at the end of a lesson or subject.	63	0	5	27	68	4	3.6	0.5
3. Determine the degree of accomplishment of a desired learning outcome at the end of a lesson.	61	0	3	33	64	4	3.6	0.5
4. Assess the quality of student learning in a class at the end of an instruction.	62	0	10	19	71	4	3.6	0.6
5. Make suggestions to students about how they develop better learning strategies.	61	0	0	39	61	4	3.6	0.4
6. Identify better learning opportunities for students in class.	62	2	2	27	69	4	3.6	0.6
7. Explore effective classroom teaching methods and strategies.	63	0	8	30	62	4	3.5	0.6
8. Guide students how to set their goals and monitor their own learning progress.	62	5	5	27	63	4	3.4	0.8
9. Make final decision about the level of learning that students achieved at the end of a lesson or subject.	61	2	3	39	56	4	3.4	0.6
10. Allow students to discover their strengths and learning difficulties in class.	62	2	11	31	56	4	3.4	0.7
11. Examine how one student performs relative to others in my class.	62	5	8	35	52	4	3.3	0.8
12. Diagnose areas for improvement of instructional activities.	63	2	5	49	44	3	3.3	0.6
13. Continuously collect learning data from students to improve instructional process.	62	0	11	47	42	3	3.3	0.6
14. Demonstrate to students how to do self-assessment.	62	3	10	50	37	3	3.2	0.7
15. Set the criteria for students to assess their own performance in class.	62	2	14	44	40	3	3.2	0.7
16. Provide examples of good self-assessment practice	63	2	9	49	40	3	3.2	0.7
17. Rank students based on their class performance to inform other school officials.	62	11	18	24	47	3	3.1	1
18. Provide information to parents about the performance of their children in school.	62	2	23	52	23	3	2.9	0.7

Teachers were given 18 statements (presented in Table 4.5) classified in five categories each referring to one type of assessment purpose namely *assessment as learning* (statement 8, 14, 15, 16), *assessment of learning* (statement 2, 3, 4, 9) and *assessment for learning* (statement 1, 10, 5). Four more statements (7, 11, 6, and 12) were also added for *assessment for instruction*, often viewed as a subset of *assessment for learning* given that its primary aim is to collect information that would be used to address inefficiencies in the teachers' instructional methods and approaches. Three other statements (13, 17, and 18) were used for *assessment to inform* which is used by teachers to obtain learning evidences for reporting purposes (Gonzales and Aliponga, 2011).

The key purpose of collecting this data was to examine if teachers' assessment purposes were in any way related to the teachers' selected forms of assessment. According to the teachers' self-report in Table 4.5, most teachers' assessment was often aimed at providing feedback in order to improve the students' learning, identifying better ways of teaching and measuring their students' gains at the end of an instructional period. The results indicate that among all the stated assessment purposes, '*providing information to parents*' was the least used (Mode=3, \bar{x} =2.98). Overall, statements referring to *assessment of learning* (statement 2, 3 & 4) and *assessment for learning* (statement 1 & 5) were among the top rated purposes of assessment. It can also be seen that item 1 on providing feedback has both the lowest standard deviation (SD=.31) and the highest mean (\bar{x} = 3.89) thus suggesting that giving feedback for learning improvement was the common purpose of most teachers' assessment.

The teachers' responses during interviews echoed the findings from the survey and suggested that teachers assessed for both formative purposes and summative. Although other purposes were put forward, the analysis of the overall interview data indicated that the primary purpose of the interviewed teachers' assessment was to gather evidence on whether their students had understood what was taught. According to the interviewees, assessment helped them to decide if they had to repeat the lesson immediately or whether they had to change how they taught the lesson. As Teacher SL clarified, assessment helped her to plan her lessons and adjust her teaching to facilitate understanding.

I want to check if the lesson has been understood well , it means, for example if I give a group work , they do it, and when there is something difficult for all students , I notice it ...I ...and I ...realise that maybe I didn't explain well. So I repeat it immediately or if time is up I make it my first focus for my future lessons. (Teacher SL)

Some teachers were specific, explaining how they used assessment tasks to see the extent to which their teaching objectives were achieved by using their pre-set achievement criteria as Teacher MR explained:

My purpose is to evaluate if they understood what I taught them. If there is something that needs to be changed or repeated [...] if they don't score at least 50% in the assessment then I know that they didn't understand well. So I look for errors on their papers. (Teacher MR)

Similarly, Teacher JS' comments reaffirmed the use of assessment as a measurement of the attainment of teaching objectives (See appendix 9):

The main purpose of every assessment is to make sure that the lesson is well delivered or not and in case the students fail the assessment you have to to... to give once again the the lesson and to make sure that they understand well. Before ...before the lesson there is a percentage the teacher wants to reach that percentage is shown by the assessment (Teacher JS)

While teachers MR and JS appeared that they used assessment to measure what their students gained and how effective their teaching was in order to take remedial measures, Teacher VN stated that his assessment was aimed at providing students with the opportunity to “express their ideas”. He also added that he used his assessment as a disciplinary measure to help students focus during the lesson.

I can even ask them when I see that they are not following well or when they say yes yes only and when I ask them they don't answer well... I ask them to take a paper and I give them a short quiz but that quiz is not marked. (Teacher VN)

Overall, the results displayed in Table 4.5 on the purpose of assessment indicate that the top three main purposes for which teachers assessed their students were (1) to provide feedback for improvement, (2) to identify better learning opportunities and (3) to measure their students' learning at the end of instruction. It was expected that teachers would report assessing students for providing feedback to students as this is the usual intended purpose of classroom assessment. It also shows that assessment of learning was among the highly rated purposes.

This was also expected given the fact that classroom assessment results were used for official reporting and for students' promotional purposes. Although teachers' self-report seemed to suggest that they used assessment for promoting learning, this may not necessarily be the formative assessment. Analysis of data from classroom observation and teacher interviews showed that the feedback provided by teachers were not always meaningful to the students (See Section 4.3).

4.2.2 Time of classroom assessment

Sub-research question 1.2 was also formulated to gather information on the specific time of instruction at which the teachers' preferred assessment forms were used in the classroom. Data was expected to provide evidence of the teachers' preferences in relation to whether they carried out their assessment at the beginning of instruction, in the middle of instruction, at the end of instruction or at any time they felt a need to do so during instruction. This information was also collected to examine whether the purposes of assessment as self-reported by teachers were also evidenced by the timing of assessment.

Table 4.6: Time of Classroom Assessment

Time of assessment	N	Never	Rarely	Sometimes	Usually	Mode	Mean	Std. deviation
		1	2	3	4			
		%	%	%	%			
1. PF* at the end of an instructional unit	66	2	3	33	62	4	3.56	0.6
2. PP* at the end of an instructional unit	60	7	0	30	63	4	3.50	0.8
3. PF at any time during the instruction	64	2	17	53	28	3	3.08	0.7
5. PP at any time during the instruction	61	2	10	72	16	3	3.03	0.5
4. PF in the middle of an instructional unit	63	3	19	51	27	3	3.02	0.7
6. PP in the middle of an instructional unit	62	3	23	57	17	3	2.89	0.7
7. PP at the beginning of an instructional unit	60	18	25	33	23	3	2.65	1
8. PF at the beginning of an instructional unit	63	11	35	30	24	2	2.62	0.9

PF*: Performance based assessment

PP*: Paper and pencil based assessment

For example, it would be expected that reporting the frequent use of assessment forms at the end of instructional unit would imply that such forms were used for summative assessment which is also usually considered as assessment of learning.

Teachers were given a list of assessment forms in two categories (performance based and paper and pencil based forms) and they were asked to indicate the specific time of instruction at which they used such forms (see appendix 5). Data on sub-research question 1.2 was also expected to provide information on the extent to which assessment forms may have been used for ‘formative assessment’, which can be done at any time during instruction unlike ‘summative assessment’ which usually takes place at the end of instruction (Black and Wiliam, 1998b; Harris, 2005). Table 4.6 provides a summary of the responses from teachers in Rwandan Lower Level of secondary schools on the types of forms of assessment that they used to assess English language learners at different times of the term. The results from data analysis show that *performance* based assessment and *paper and pencil* based assessment were both given at the end of instructional unit. On the other hand, the results highlighted in Table 4.6 suggest that both ‘*performance*’ and ‘*paper and pencil*’ forms were less frequently used at the beginning of instructional unit as illustrated by the lowest mean ($\bar{x}=2.62$) for paper and pencil and ($\bar{x}=2.65$) for performance assessment respectively. It can also be observed that ‘*performance*’ based forms were more used to assess students in the middle of instructional unit than ‘*paper and pencil*’ based assessment ($\bar{x}=3.02$ and $\bar{x}= 2.89$ respectively).

Analysis of teachers’ comments made during interviews reveals some discrepancies with the results from the questionnaire survey regarding the extent to which the ‘*performance*’ assessment was used in the classroom. While data in Table 4.6 suggests that ‘*performance*’ based forms were nearly as frequently used as ‘*paper and pencil*’ based forms, data from the teachers’ interviews suggests that ‘*performance*’ assessment was rarely used in day-to-day classroom assessment. It was also found that performance assessment was often used at the end of instructional units partly because their preparation was reportedly time consuming and that their availability was very limited as Teacher MR stressed:

It’s usually written assessment [...] oral is given for example after one month of preparation after students finish reading books...it’s time consuming... A week may end without oral assessment. Written assessment is the one mostly used. It may

be used twice or three times a week or so. Oral ...I prepare it for example I can decide that I will have two oral assessments in a term... But they are very few.
(Teacher MR)

Data from classroom observation indicated that teachers tended to give written assessment activities (see Table 4.4) during instruction. Examples of these assessment activities were recorded in 4 out of 7 observed classrooms and no performance assessment was given in any of these classrooms. It was noted that the teachers' oral questions at the beginning and throughout instruction was the only non-written assessment commonly used by all the observed teachers. Such oral questions apparently aimed to recap on the previous lessons when used at the beginning, or for calling the students' attention during instruction. These results were not far from the expected. As the teachers' assessment was perceived to be summative oriented, it was also expected that most assessment forms would be used at the end of instruction to generate evidence of learning achievement for reporting purposes. Comments from some teachers during interviews and data from classroom observation also seemed to suggest that unlike 'paper and pencil' based forms that seemed to be used at any time of instruction, 'performance' based forms were more likely to be used at the end of certain instructional unit.

4.2.3 Content focus of assessment

Data was gathered through questionnaire survey to answer part of sub-research question 1.3 regarding the content focus of teachers' classroom assessment. This information on the content of assessment was collected to help generate a holistic understanding of the teachers' assessment preferences in relation to the assessment of the four language skills (i.e. speaking, writing, reading and listening) and language systems of grammar and pronunciation. Data on how frequently the receptive skills were assessed was sought to facilitate a comparison with productive skills and a better comprehension of the extent to which classroom assessment was focused on English language skills in general. In addition, information on non-achievement factors was also gathered to determine the extent to which teachers attached importance on factors such as student's attendance and student's effort. Non-achievement factors have been found to affect the students' self-efficacy (Alkharusi et al. 2014) and data on how teachers dealt with these factors was expected to help with the interpretation of some results on students' self-efficacy.

In the current study, teachers were asked to indicate how frequently their assessment was centred on English language skills and systems or on non-academic achievement factors within a limited time of one week. A one week time frame was used to identify the extent to which teachers focused on different contents. A Likert scale with specific measurements points were used to elicit specific information indicating precise frequencies: ‘*More than three times a week*’, ‘*Two to three times a week*’, ‘*Once a week*’ and ‘*Once every two weeks*’. As highlighted in table 4.7, the teachers’ assessment was mostly focused on grammar as was reported by the majority of respondents (67%) and demonstrated by highest mean of 3.57 on a 4 point scale. The lowest standard deviation (SD=.66) is also an indication that all teachers agreed that grammar was the main focus of their assessment. The non-achievement factors were also frequently assessed as highlighted by the second highest mean (\bar{x} =3.55) for student’s ‘class participation’, and \bar{x} =3.50 for student’s ‘class attendance’.

Table 4.7: Content focus of teachers' assessment

Focus of assessment		N	Once every two weeks	Once a week	Two to three times a week	More than three times a week	Mode	Mean	Std. deviation
			1	2	3	4			
			%	%	%	%			
1	Grammar	63	0	9	24	67	4	3.57	0.6
2	Student class participation	64	3	13	11	73	4	3.55	0.8
3	Student class attendance	64	8	6	14	72	4	3.5	0.9
4	Writing	64	8	16	20	56	4	3.25	0.9
5	Speaking	60	7	20	23	50	4	3.17	0.9
6	Vocabulary	63	6	16	35	43	4	3.1	0.9
7	Pronunciation	61	8	23	21	48	4	3	1
8	Student overall effort	63	13	11	33	43	4	3	1
9	Reading	63	8	28	29	35	4	2.9	0.9
10	Listening	57	39	17	12	32	1	2.37	1.2

Data on teachers' assessment focus also suggested that overall, teachers' assessment was inclusive (Mode=4). However, it can be seen in Table 4.7 that productive skills were more assessed than receptive skills. These results show that speaking and writing had higher means ($\bar{x}=3.17$ and $\bar{x}=3.25$ respectively) while Listening remained the least assessed skill ($\bar{x}=2.37$).

These results seem to agree with the results from classroom observation which confirm that the teachers' classroom assessment was often based on grammar (see Table 4.4). The analysis of the recorded videos indicate that assessment in five of the seven observed classrooms (71%) was based on grammar while only two teachers (29% of the observed classrooms) focused on paraphrasing techniques and summary writing. None of the observed teachers gave speaking or listening tasks. Although Teacher VN asked his students to read a text that they were reportedly going to summarise at some point in their course, the reading activity was quick and brief. However, I recognize that the current data was gathered in a relatively short time and one might have expected to see more assessment tasks should more lessons have been recorded for a longer period. Nonetheless, based on my own past experience as an English language learner and English language teacher in Rwandan secondary schools, the lack of oral skills assessment was not unexpected. As was anticipated, teachers used a great part of their time explaining grammar structures and writing the rules on the chalkboard which has been the tradition of many teachers in Rwanda (Odeke, 2015; Sibomana, 2014). Data in the current study also showed that oral display questions to individual students were often centred on grammar and teachers hardly asked referential questions.

Evidence of the teachers' emphasis on assessment of grammar and little attention to the assessment of language skills was also found in other data sets. It was noted for example that results from classroom observation showed that teachers always used 'paper and pencil' forms to assess their students on grammar related material. In addition, the teachers' comments during interviews confirmed that their assessment was often focused on grammar and that the 'paper and pencil' assessment was usually used in their classrooms. This seems to imply that the limited assessment of other language skills such as speaking could explain why performance assessment such debates and oral presentations were also very rarely used. By focusing on grammar, teachers were also likely to use controlled assessment forms such as short sentence

completion, fill in gaps or multiple answer questions. Assessment focused on content other than on the practice of the language skills may thus influence the use of paper and pencil based assessment forms.

4.2.4 Sources of assessment tasks

Part of sub-research question 1.3 aimed to explore the sources of assessment tasks that teachers used in their everyday classroom assessment. The gathered data was expected to provide information such as whether assessment items were developed by teachers themselves or in collaboration with their colleagues. Some activities may also be downloaded from internet or taken from external papers. They may also be embedded in textbooks that teachers use in their everyday teaching. On a 4-point Likert scale using *usually* =4, *Sometimes*=3, *Rarely* =2 and *Never*=1, the majority of teachers (66%) indicated that they usually used assessment tasks that they developed on their own (\bar{x} =3.65). The results in Table 4.8 also indicate that teachers often used *textbook-embedded* question items as well as *external items* from past exam papers. Teachers also rarely used item questions developed together with their colleagues and rarely used to internet materials for their classroom assessment.

Table 4.8: Source of teachers' assessment activities

Source	N	Never	Rarely	Sometimes	Usually	Mode	Mean	Std. deviation
		1	2	3	4			
		%	%	%	%			
1. Question items developed by myself	62	0	2	32	66	4	3.6	0.5
2. Textbook-embedded items	63	3	2	40	55	4	3.4	0.7
3. Items from external exam papers	63	0	13	65	22	3	3.1	0.5
4. Prepared together with other teachers	63	18	30	35	17	3	2.5	0.9
5. Items found on the Internet	62	23	32	37	8	3	2.3	0.9

Data from interviews with teachers revealed that, in addition to the use of textbook-embedded tasks, teachers used assessment tasks from many other sources. In their responses,

the majority of teachers reported that they used tasks that they developed on their own and tasks from textbooks. They pointed out that textbook embedded tasks were commonly used because they were easily accessible from different places including mainly from school libraries. Nonetheless, Teachers SF, SL and MR pointed out that they used their own notebooks from high school. Teacher SF explained that as he taught the same content as what he studied at high school and given the shortage of books, he used notes that he had taken during his high school studies and followed the same approach that his teachers used.

[...]This school, especially ours which is a Twelve Year Basic Education, they are poor in terms of resources. And where do I found these in case of ...these ...assessment or the tasks? First we have the books, some little little books [...] there is a programme and the programme also they are telling you this and this and this, and there are also personal books, what you have used maybe in your study...Teacher's personal books ... in some cases, what you are teaching is what you have studied also, you know ...it is like that. (Teacher SF)

Teachers teaching in 3rd forms also stated that they looked for past papers of the National Ordinary Level Examination to help their students prepare that same examination. In addition, some teachers added that they used internet materials where it was available as explained by Teacher VN in the extract below.

There is a library there... I take books sometimes... I take books other times I take national examinations of previous year for the third form but sometimes I consult internet because here there are some lessons that are not found in the books that they have brought. (Teacher VN)

Teacher VN indicated that internet was used as a second option when their textbooks did not contain lessons that they wanted to teach. The use of internet material was not reported by other teachers. This was not surprising given that access to internet connected computers is still very limited at many secondary schools. On the other hand, the teachers' reported use of assessment materials from their previous grammar based language classes was not anticipated. This explained why some teachers still stuck on the grammar based assessment in practice although they were in support of the use of performance assessment. Having been taught and assessed on grammar, many teachers seemed to use the same approach. The focus on grammar was also partly encouraged by the nature of the textbooks available at some schools. The

analysis of interviews and observation data revealed that teachers who often relied on textbook embedded tasks used paper and pencil assessment forms to assess grammar.

4.3 Teachers' methods of providing feedback

Having presented the results on the forms of assessment that teachers used in their classrooms, this section is a report of the results on the types of feedback that was used in the classroom. Sub-research question 1.4 (See Figure 3.1) was formulated to collect information on the different methods that teachers used to provide feedback to their students after assessment tasks. A four point Likert scale was used to record the teachers responses on how frequently each type of feedback was used in their classroom using *Usually* =4, *Sometimes*=3, *Rarely* =2 and *Never*=1. Information on the nature and frequency of feedback that teachers provided was key to this study as feedback can affect the students' self-efficacy through teachers' approval of students' work and praise for good performance which inform learners about their learning strengths (Beleghezadeh and Masoun, 2013; Thomas, Usher, and Mamaril, 2012).

Table 4.9: The methods through which students received feedback

Methods	N	Never	Rarely	Sometimes	Usually	Mode	Mean	Std. deviation
		1	2	3	4			
		%	%	%	%			
1. Collective verbal teacher feedback	61	2	3	28	67	4	3.6	0.6
2. Written collective teacher feedback	61	0	7	36	57	4	3.5	0.6
3. Student peer feedback	59	4	8	41	47	4	3.3	0.7
4. Individualized verbal teacher feedback	62	3	18	37	42	4	3.1	0.8
5. Individualized written teacher feedback	63	3	18	38	41	4	3.1	0.8
6. Test results only (marks)	63	24	2	36	38	4	2.8	1.1

One question item in the questionnaire survey was also used to investigate whether, in addition to collective or individualized teacher feedback, teachers also encouraged students to provide feedback to each other (peer feedback). On a cut-off point of 4, the highest mean (\bar{x} =3.61) in Table 4.9 indicates that teachers mostly used collective verbal feedback. During

classroom observation, the tendency to use collective feedback for all types of assessment was a common feature for the majority of teachers who were observed. Verbal collective feedback was often provided in the form of comments, usually after the distribution of individual test results to students.

Table 4.9 also shows that written collective teacher feedback was frequently used ($\bar{x}=3.51$) with the lowest standard deviation ($SD=0.6$). It was noted during classroom observation that collective teacher feedback often consisted of comments or answers to assessment questions written on the blackboard and usually recorded by students in their notebooks. The survey data also demonstrated that test results in the form of marks were rarely used alone without further comments ($\bar{x}=2.89$). However, it can also be observed that 'test results only' type of feedback has the highest standard deviation ($SD=1.1$) which seems to show the teachers' divergent views on this question. Overall, it is clear from Table 4.9 that collective teacher feedback was more frequently used than individualized feedback both at the oral and written feedback levels. This was also reflected in the results from classroom observation.

In addition to feedback provided to individual groups, the analysis of data from classroom observation show that teachers often provided feedback collectively by writing some explanations on the chalkboard. Some teachers also asked students to take turns to write the answers on the blackboard. As shown in Table 4.4, Teacher PL assessed his students on grammar and the correction was done collectively on the chalkboard. Teacher PL also distributed papers to different pairs and asked them to mark each other's copies. Although students thought that the marks would be officially recorded, he explained to me during one to one interview that the marks would not be recorded. He explained that it was done so that students could take this task more seriously for the next time. Teacher PL's peer assessment was not used in any of other observed classrooms. However, comments from some interviewed teachers suggested that they used peer assessment only to help students focus and would not record the marks.

During the interview, the teachers indicated that they used diverse methods to provide feedback although the classroom observation suggested that oral marks and collective feedback were the most frequent. Interviewed teacher explained that their feedbacks were provided both

verbally and/or in written forms. They also stressed that both individual and collective feedback were provided in the classroom. These could be given orally by reacting to the student's answer with praise or by asking the rest of the class to clap for the right answer. Teacher JS explained that she asked her students to clap for the right answer because she considered it as an “encouragement”. In some cases, teachers asked follow up questions or designated another student to answer the same question when no satisfactory answer was given. Oral feedback was also provided in the form of advice to failing students. Teacher MR stated that she preferred to use oral advice to written comments to avoid possible students' misinterpretation of the feedback as one of her colleagues had once experienced.

I ask him or her to come and see me so that I can advise them on their failure. This is individual because I want to take such an opportunity to discuss his or her behaviour and how it is affecting them. Otherwise I tell them collectively...It's not written because I know about a bad experience at my previous school, many students at that school accused the teacher of harassment. I've had fear of doing it since then. I never write a comment on the student's work. Only sometimes I put a question mark. When I need to talk to them I call them. (Teacher MR)

Teacher MR's reason for not providing written feedback to the failing students was unexpected. On the one hand, students who consistently failed the assessment could not be given individualised written feedback. However, as Teacher MR reported, failing students could get verbal advice only while high performing students could get written feedback. This feedback was reportedly intended to encourage the student on keeping up the good work by writing on the student's paper phrases such as “keep it up” “congratulations” or “excellent”. Teacher MR's approach of providing oral feedback in the case of students' failure was not reported by any other teacher and it seemed to be born out of her previous experience. However, all other interviewed teachers acknowledged that they used the same method for positive feedback such as writing words of praise and encouragement on the students' answer sheet or asking the whole class to clap for the correct answer. While Teacher MR's preference was oral feedback to written comments, other teachers (e.g Teacher VN and SF) indicated that they used both oral and written feedback.

All teachers pointed out that they used collective feedback that they usually wrote on the chalkboard. They explained that such feedback could simply be answers to assessment

questions (e.g Teacher PL as shown in Table 4.4) or extended explanation of a topic covered during assessment activities. Teachers often asked students to write answers on the blackboard and in some cases, the teacher targeted weak students and asked them to answer.

Feedback is on the chalkboard, the correction of the questions. If it is a work for example, we correct on the board and I ask a student that I know is weak, to go to the blackboard to write the answer. When I know that the student knows the answer, I don't send him/her. I send his or her classmate to write the answer on the board and when it is an incorrect answer, others say no and they go and correct him or her. (Teacher PL)

Teacher PL's method of involving weak students in the classroom activities was observed in other classrooms. Some teachers called out the names of students and asked them to write their answers on the chalkboard. The teachers seemed to believe that it was the best way of helping the struggling students although the singled out students appeared anxious especially when they failed to give correct answers. Teacher PL in the extract above appeared to suggest that the student benefited from being corrected by colleagues on the blackboard. However, there was no convincing evidence that weak students benefited from this practice as some students may see their weakness being exposed. In another classroom, a student who was asked to write the answer on the blackboard by Teacher FR became anxious and was left confused after she wrote a wrong answer. Teacher FR simply asked another student to write the correct answer and no effort was made by the teacher to check the understanding of that weak student. Overall, this teacher's practice of targeting weak students for answers during oral question or during correction of assessment tasks was noted at some point during lessons in all the seven observed classrooms. This could be observed mainly during oral questions where teachers often targeted students who did not raise their hands to answer.

In addition to written collective feedback that was common to all teachers, marks were also mentioned by all teachers as another form of written feedback. This was often individual. Teachers stated that after marking each student's work, marks were communicated by returning and distributing the marked papers, after which correction on the board often followed. Individualised comments were not as frequently used as marks or oral feedback. However, teacher VN and SF confirmed that they always accompanied marks with individual comments.

SF: I write on these papers when I give feedback, written feedback [...] I give a grade and this grade is accompanied by the comments: 'for next time it would better when you try to read more books because you are poor in vocabulary, it is time to go in the library and you borrow the books, for example this one...'

Researcher: And you do that for every student?

SF: exactly

Researcher: All the time?

SF: Exactly. When I give a written one

Teacher VN provided reasons why he always provided individual written feedback, citing distractions as possible setback when feedback is given orally and collectively or when answers are simply written on the blackboard.

Normally I give most of time I give oral feedback to the whole class ... but sometimes I take a paper where there is an empty space on their paper, I write some words, if I see that I can write... 'Use capital letter at the beginning of every sentence' at that paper [...] I choose that way because when I give oral feedback in front of the whole class ... some are following but others are not following that is why those comments... I write those comments under or above marks so that if they are going to see the mark they also see... (what's going on). (Teacher VN)

On the other hand the students' responses during focus groups reemphasized the use of collective feedback particularly correction done on the chalkboard. They confirmed that marks were the usual sort of feedback they got. As highlighted in the following extract from Group C below, the students also reemphasized that teachers used collective feedback provided in the form of correction done on the blackboard.

S5: He marks and returns the marked papers. He comes and corrects the questions on the chalkboard and we check our answers.

S1: It means he corrects on the blackboard first, and then you see what you didn't do well. After the teachers finishes to mark our papers, he brings them back and show us marks. If there is an error in his marking, then you can claim your marks. For example if he forgot to tick a correct answer.

A typical cycle seemed to be common to all teachers: first teachers collected papers, marked them and returned them for students to see their marks. This was often followed by collective correction on the blackboard after which students checked whether there were no errors in the teachers marking. This was the same process for both group and individual works.

Questions on teachers' methods of providing feedback also investigated the preference of teachers in relation to assessing and providing feedback to individual students or in groups. Appropriate question items were used to elicit information on the teachers' use of individual, pair or group works, whether peer assessment and peer feedback was used as well as on how teachers monitored their students during assessment. The interview data indicated that the majority of teachers preferred pair and group work to individual work for different reasons. Some of the cited reasons included that group work was not as time consuming as individual work because teachers were able to provide feedback by talking to and/or marking the students in groups instead of having to deal with one student at a time. It was also mentioned that pair or group works helped students to explain and provide feedback to each other. Teacher SF justified his preference of group work to individual work, stating that the former was recommended in the new National English Curriculum. He added that students were "teachers to each other", which, according to Teacher SF, meant that students learned from each other during topic discussions in groups or pairs.

SF:...and you know in competence-based curriculum ...

Researcher: the new one

SF: yeah the new one. They are saying that as long as you have 30 students in class and you... teacher + ...you are 31. You have to have more than..., you have to have 31 teachers also.

Researcher: ah sorry! What does that mean?

SF: As long as you have 30 students in a class, and you teacher plus this one 30, it means plus 1 teacher, it means there are 31 people... Finally, you are 31 teachers.

However, some teachers stated that they preferred individual work to group work for most assessment. Teacher VN explained that he preferred using individual work because he wanted to get his students used to working individually as this was the method used in the national examinations. Teacher JS agreed with Teacher VN stressing that using group works meant that there was a possibility that students could be distracted and not focus on the task which would be a big waste of time. In general, teachers stated that they preferred to use group and pair works for unmarked assessment. With regard to peer-assessment, the majority of teachers pointed out that they used them rarely and cautiously because it could cause problems including students' confusion or even conflict if such assessment was marked. Teachers indicated that

they used peer-assessment only for unmarked assessment or for some multiple choice tasks where answers were clear. One unexpected answer was provided by Teacher SL who referred to the ‘history of Rwanda’ to justify why she never used peer assessment. She explained that she never used peer-assessment over fear that she might pair up two students whose families were not good neighbours.

Researcher: What about peer assessment? Asking students to assess and mark each other’s work.

SL: I don’t do it. Because seeing our past, the history of Rwanda, these are day school students who go back home. I may not know that they are neighbours and not on good terms. If I give them each other’s paper and they mark it unfairly, then it can create misunderstanding ...

Researcher: That is when you give them home work. Let’s focus on what they do in the classroom now.

SL: Even in the classroom... I keep that in mind. I can’t ask them to assess each other,

Researcher: So you don’t think they can do it well?

SL: They may do it well but I really don’t know them well.

Teacher SL seemed to refer to the 1994 genocide. This was an unexpected comment given that no cases of social tensions being reflected in the classroom were reported in Rwandan research before. However, it may be true that the scars of genocide against Tutsis in 1994 have not yet faded and incidences of the kind may be observed in some schools. Moreover, the fact that no other teacher mentioned this issue may not necessarily mean that it was not present in their classrooms. Teachers, like many other Rwandans, may find it uncomfortable to discuss issues relating to the genocide and its effects today and reference to this sensitive topic may be avoided by researchers, hence the lack of grounded evidence to support the impact that the genocide has had on Rwandan teacher classroom practices. In addition, the teacher interview sample may not be large enough to allow the emergence of many diverse ideas that would be suitable for the generalization of the results. However, although it may be realistic to say at this point that Teacher SL’s claims of the effects of social tensions on her classroom practices constituted a one-off exception, it is also fair to recognise that due to Rwanda’s 1994 tragedy of genocide against the Tutsis, some teachers like SL may be overly sensitive to accusations of stirring up social conflict by engaging students in activities that may bring up memories of the past.

In sum, the results of the analysis showed that teachers used different methods of providing feedback for motives and purposes that differed from teacher to teacher. Some selected their methods based on their past experience, their school context and their educational background. Overall, the teachers seemed to be convinced that their methods of providing feedback were effective in helping their students in their learning despite the fact that some of their methods could be off-putting especially for the struggling students. It was also found that teachers preferred to use collective feedback, most of which was written and marks based. Data from classroom observation suggest that most teachers' feedback was not necessarily constructive as it provided little or no information on the students' weaknesses, strengths or ways for improvement for the struggling students. However, it is important to note that the small sample size for the classroom observation was a limitation and not enough data was collected for drawing solid conclusions.

4.4 Students' perceptions of teachers' assessment practices

Sub-research question 1.5 was formulated to gather data on the students' perceptions of assessment practices used by their teachers of English. The students' perceptions of teachers' classroom assessment were investigated to examine their relationship with the students' self-efficacy (See Section 5.2). Some research studies have indicated that students' perceptions of teachers' classroom assessment practices are highly related to the students' self-efficacy (Alkharusi, 2013; Alkharusi et al., 2014; Dorman et al., 2006). As indicated in Table 4.10, eight statements were provided in part two of the students' questionnaire to collect information on the students' views on the teachers' classroom assessment practices including feedback. A 4-point scale was used where the respondents were asked to choose one from *Strongly Agree=4*, *Agree=3*, *Disagree=2*, or *Strongly Disagree=1*

The students were only asked to rate how satisfied they were with their teacher's feedback in general. More follow up questions were asked during focus groups to know more about the nature of the feedback provided by teachers. Students were asked to indicate the extent to which they agreed with statements regarding (1) the extent to which the assessment matched the instructional content (congruence), (2) the extent to which the assessment tasks

reflected the real life language use (Authenticity), (3) How teachers involved their students in classroom assessment decisions (Student consultation), (4) the transparency of assessment aims (Transparency), (5) how diversified the assessment tasks were (Diversity), (6) whether feedback was always provided (Feedback), (7) if the students thought that their teachers were fair in marking (Fairness of marking) and (8) how frequently they were assessed in the classroom (Frequency of assessment).

Table 4.10: The students' perceptions of the teachers' classroom assessment practices

Statement	N	Strongly Disagree	Disagree	Agree	Strongly Agree	Mode	Mean	Std. deviation
		1	2	3	4			
		%	%	%	%			
1. I think the teacher always grades my work fairly.	1254	3.3	7.1	41.1	48.4	4	3.3	0.7
2. The content of classroom assessment matches the content of instruction	1246	1.4	3.4	55.2	40	3	3.3	0.6
3. I think the current classroom assessment prepares me for better use of English in the future.	1228	2.5	7.6	50.3	39.6	3	3.2	0.7
4. The teacher uses different kinds of assessment activities (example: written tests, oral presentations, etc).	1244	3.1	8.9	50.5	37.5	3	3.2	0.7
5. Overall, the teacher gives us enough assessment activities.	1253	3.1	7.7	46.9	42.2	3	3.2	0.7
6. I always get enough feedback from my teacher after assessment tasks.	1244	5.1	17	46.6	31.3	3	3.0	0.8
7. It is always clear what the teacher wants me to do on assessment activity.	1247	3	21	56.7	19.2	3	2.9	0.7
8. I take part in deciding what needs to be on my assessment activity and how to do it.	1244	33.4	40.4	21.1	5.2	2	1.9	0.8

The results in Table 4.10 show that the respondents mostly agreed that teachers were fair in marking (\bar{x} = 3.35 with the highest mode =4) and that assessment was aligned with instruction (\bar{x} = 3.30). It can also be observed that *congruence* of assessment tasks has the lowest standard deviation (SD=0.6) indicating homogeneity in the respondents' answers on this

item. The two lowest means are for *students consultation* ($\bar{x}=1.98$) and *transparency* of assessment tasks ($\bar{x}= 2.92$). The students' answers on how they perceived their teachers' assessment practices may not necessarily reflect the reality of what happens in the classroom. For example, the students' answers on the *diversity* and *sufficiency* of assessment activities suggested that teachers provided sufficient and diversified assessment tasks. However, the results from interviews and classroom observation seem to indicate that only a narrow range of assessment forms and assessment activities were used in the classroom. Nonetheless, the results displayed in Table 4.10 provide useful information that helps to compare the students' perceptions for different teachers' assessment practices.

During interviews, questions on the students' perceptions were similar to questions used in the questionnaire survey. Participants were asked to state how they viewed their classroom assessment; whether they thought it was aligned with what they studied in the classroom, their perceived level of task difficulty or whether they were satisfied with the frequency of assessment. As highlighted by student S1, S2, and S3 from Group A, assessment tasks were usually aligned with what was taught in the classroom. The students also pointed out that task difficulty depended on how well they made effort in reviewing their lessons. This was important because it could imply that the students understood well that their effort could be a factor influencing their perceptions of classroom assessment and their perceived self-efficacy.

S1: it's not difficult

S2: Because it's usually about what we studied

S3: It also often depends on how hard you have studied and revised notes. The way you revise what the teacher taught you.

Students indicated that doing assessment aligned with what they studied in classroom reduced task difficulty. They also stated that they were used to some format such as questions starting with the question word "what". The students added that changing the format of the questions could lead to confusion and misunderstanding of what they are required to do. As can be inferred in what students S1 and S2 from Group B said below, the students' performance on assessment could depend on their familiarity with the methods used which may increase or decrease the level of task difficulty.

S1: In general, assessment tasks are easy. However, sometimes it is difficult to understand what needs to be done. You may be able to read the text but fail to understand it.

S2: When the teacher changes the method of assessment which is familiar to us, it becomes difficult to us. We are only familiar with questions starting with “what is...”

Student S2 response above that a change in question structure during assessment would become a challenge could be indicative of the adverse impact that the lack of diversity in assessment forms had on the students. In addition, S2 claim seems to contradict the results from the student questionnaire survey (see Table 4.10) which suggested that the students were satisfied with how teachers used diverse assessment forms. This indicates some limitations with the students’ reported perceptions as some may not necessarily reflect the reality of the teachers’ practices. However, there seemed to be similarities in the findings from interviews and classroom observations data which tends to show that teachers used a very limited range of assessment forms with a particular focus on controlled assessment as was implied in S2 answer.

4.5 Summary of the chapter

The presentation of results in chapter four has been focused on the results from the analysis of data on the teachers’ assessment practices mainly the forms of assessment that teachers of English used in the classroom and the methods of providing feedback to their students. Chapter four also presents the results on the purpose and time of classroom assessment, the content focus of assessment as well as the source of assessment tasks. The chapter also reports on the students’ perceptions of the teachers’ classroom assessment practices as a whole. Overall, the results obtained from the analysis of the quantitative data from teachers point to the use of both performance assessment forms and paper and pencil assessment forms. Teachers also seemed to report during the questionnaire survey that they assessed mainly for formative purposes, with the primary objective to provide constructive feedback during instruction (See table 4.5). However, the results from the analysis of qualitative data from the students’ focus groups and classroom observation suggest that paper and pencil based forms of assessment were the most frequently used. The use of summative assessment where marks and grades were used to provide students with feedback was also found common among teachers. The thematic

analysis of data from the students' focus groups and the teachers' interview (See appendix 11) also confirmed that teachers often used the summative assessment usually at the end of instructional units. Based on my previous experience as a teacher in the Rwandan context, the prevalent use of summative assessment and high reliance on paper and pencil forms were expected. However, the mismatch between some results from the two sets of data (qualitative and quantitative) was unpredicted and may be a result of factors such as social desirability bias (See section 3.3.1). The results also show that the students had high positive perceptions of the teachers' assessment practices in spite of the teachers' limited use of diversified assessment forms and overall prevalence of summative assessment. This could be explained by the fact that students seemed to have limited knowledge of the different forms of assessment and methods of feedback that could be used to maximise their learning potential.

Chapter 5: Students' self-efficacy and its relationship with teachers' assessment practices

5.0 Introduction

So far the presentation of results has focused on the practices of the teachers in relation to classroom assessment. The presented results were based on data collected from the teachers, the students and from classroom observations. Specifically, Chapter four has provided a descriptive account of what teachers reportedly did in relation to their classroom assessment mainly within the scope of assessment forms, use of rubrics and teachers' methods of providing feedback. The results show that *paper and pencil* based forms of assessment were predominantly used by teachers for their classroom assessment and no reports of rubrics use were recorded. It was found that the teachers' assessment was also usually summative oriented and mainly focused on grammar. Verbal and collective feedback provided to whole class was the main methods through which the students got feedback from the teachers. The use of marks as a way of providing feedback to students was also found common among teachers. Chapter four also reports on the results obtained from the students on how they perceived their teachers' classroom assessment practices in terms of assessment congruence with instructional content, transparency of assessment aims, authenticity and diversity of assessment tasks, frequency of assessment, student consultation as well as teacher feedback and fairness in marking (see Section 4.3). Overall, the students reported positive perceptions of the teachers' practices. However, the students' perceptions were reported as low in relation to transparency of assessment aims and student consultation (see Table 4.10).

The current chapter focuses on data about the students' self-efficacy in the four English language skills and examines the extent to which the teachers' assessment practices and the students' perceptions described in chapter four were both related to the students' self-efficacy. The chapter presents the results of the analysis of data obtained from the students' questionnaire survey and the students' focus group interviews regarding their self-efficacy for

speaking, reading, listening and writing in English. The presented findings pertain to research question two: *To what extent is the students' self-efficacy in the four English language skills related to the teachers' assessment practices?* The presented results specifically answer the three sub-questions (1) *What is the level of the students' self-efficacy in the four English language skills?* (2) *How is the students' self-efficacy in the four English language skills related to the students' perceptions of the teachers' assessment practices?* and (3) *How is the students' self-efficacy in the four English language skills related to the teachers' assessment practices?* Each section reports on data obtained to answer each sub-research question and as in chapter 4, the findings from the quantitative data are presented first and supported by the qualitative results derived from the students' comments during their interviews.

5.1 Students' self-efficacy in the four English language skills

Turning now to the students' self-efficacy, this section presents the results of the analysis of data from the student questionnaire survey regarding the students' self-efficacy and from the students' focus group interviews. It reports data on sub-research question 2.1 which was asked to collect information on the extent to which the students felt capable of successfully completing tasks in both productive (Speaking and writing) and receptive skills (reading and listening) in English. Obtaining data on their self-efficacy in the four language skills was expected to help identify the skills in which the students felt most self-efficacious and later to identify the extent to which the students' self-efficacy in each skill was related to the teachers' assessment practices in the classroom. It was hypothesized that a relationship existed between the teachers' assessment practices, specifically between the use of assessment forms and types of feedback used in the classroom, and the student's self-efficacy for speaking, reading, listening or writing in English. This is based on the Social Cognitive Theory (see Section 2.5) by Bandura (1986) according to which having success in executing tasks boosts student's feelings about their ability. In connection to this therefore, the current study first gathered information about the students' self-efficacy in order to compare it to the teachers' assessment practices and to explore the relationship between these two variables.

During the questionnaire survey, students were asked to indicate how certain they were that they could effectively perform each stated tasks in English by choosing a point on a 100

point scale, i.e. from 0 (*Not certain at all*) to 100 (*completely certain*). The measurement was adapted from Bandura's (2006) 11-point scale where statements were changed to relate them to the topic of the research (See section 3.4). Statements corresponding to the tasks were selected from the National English Language Curriculum for Rwandan lower level of secondary schools (MINEDUC, 2011). The internal consistency for the subscale of each skill was calculated and high Cronbach's alpha coefficients indicated a high level of internal consistency for speaking ($\alpha=.91$), for writing ($\alpha=.93$), for reading ($\alpha=.92$) and for listening ($\alpha=.94$). The whole scale for the students' self-efficacy was found to be highly reliable (20 items, $\alpha=.97$). A scale with a Cronbach's alpha of .60 or higher is considered as reliable (Cohen et al., 2007).

The first set of analysis was exploratory, to measure the normality of data distribution in relation to the students' self-efficacy in each skill. The test of normality can help the researcher to know more about the characteristics of data and to identify the appropriate statistical test to use between parametric and non-parametric tests (Dornyei, 2007). In this study, the normality of data with regard to the students' self-efficacy was measured using the Kolmogorov-Smirnov (K-S) test and Shapiro-Wilk test. K-S and Shapiro tests results showing a non-significant test ($p>0.05$) indicate a normally distributed data (Field, 2009).

Table 5.1: Results from the normality test of the students' self-efficacy for the four skills

	Kolmogorov-Smirnov ^a			Shapiro-Wilk		
	Statistic	df	Sig.	Statistic	df	Sig.
Speaking	.081	1245	.000	.945	1245	.000
Reading	.083	1245	.000	.947	1245	.000
Writing	.061	1245	.000	.962	1245	.000
Listening	.051	1245	.000	.973	1245	.000

a. Lilliefors Significance Correction

The results presented in Table 5.1 show that both Shapiro and the K-S tests on the student' self-efficacy are significant ($p=.000$) for all four English language skills. This means that the students' scores on self-efficacy in Speaking $D(1245) = .08, p=.001$; in Reading $D(1245) = .08, p=.001$; in Writing $D(1245) = .06, p=.001$, and in Listening $D(1245) = .05, p=.001$ were all non-normally distributed. In addition, the values of skew and kurtosis were also calculated to back up the K-S results on the non normality of the data. The values of skewness and

Kurtosis were converted into z -scores and compared to 2.58 threshold which is recommended in large samples. The z -score value greater than 1.96 indicates a significant test which means that data is significantly different from normal distribution (Field, 2009). The calculated z -score in the current study was found to be greater than 2.58 for all the four skills. Following these results from both sets of tests, it appeared that the normality assumption required for parametric tests was violated. The non parametric tests were thus adopted for further analysis relating to the students' self-efficacy. Due to the observed skewness of data, the mean was used in conjunction with the median and standard deviation for results comparison purposes. In addition, instead of using the independent T-Tests, the Mann-Whitney U test was found appropriate to test the significance of differences in the students' self-efficacy on two group variables. Similarly, the Kruskal-Wallis test was used in place of ANOVA to test the significance if differences in the students' self-efficacy based on a more than two group variable.

Table 5.2: Central tendency of the students' self-efficacy for each skill

	SE for Speaking	SE for Writing	SE for Reading	SE for Listening
N	1251	1248	1249	1245
Median	75	67	74	63
Mean	71.2	64.2	69.8	61.2
Std. Deviation	20.61979	23.18132	21.83037	23.70080

As can be seen in Table 5.2, the test of central tendency of the students' self-efficacy (median) suggests that students felt more self-efficacious in speaking and reading than in writing and listening. Speaking and reading also have the lower standard deviations of $SD=20.61$ and $SD=21.83$ respectively indicating a lower degree of data dispersion in these skills. On the other hand, the results indicate that students felt least self-efficacious in listening as demonstrated by the lowest median score (63). These results were expected at this point given the fact that listening was not among the top assessed skills in the classroom (see table 4.7). In the context of this study, less assessment focus on a given skill would suggest that students would not feel self-efficacious due to the lack of opportunities to experience success or witness their colleagues performing well in that skill. Some respondents may as well

overestimate their self-efficacy (Pajares, 1996; Artino, 2012) due to their limited awareness of their true abilities to perform tasks in a given skill.

Further analysis was carried out to explore possible differences in the students' self-efficacy based on factors such as the student's gender and age. To measure the statistical significance of these differences, appropriate statistical tests were applied to different sets of data. The analysis of differences in the students' self-efficacy levels was focused on the students' background and demographic factors. The Mann-Whitney U test was used to examine differences in the students' self-efficacy based on gender and school learning mode whereas the Kruskal-Wallis test was used to identify the extent to which other students' background and demographic factors such as age, school learning mode and parents' education levels affected the students' self-efficacy. As illustrated in Table 5.3, the Mann-Witney mean ranks suggest that the students' self-efficacy varied as a function of gender whereby male students had higher levels of self-efficacy in all English language skills than their female classmates.

Table 5.3: Mann-Whitney Mean Ranks of the students' self-efficacy by gender

	Gender	N	Mean Rank	Sum of Ranks
Self-efficacy for Speaking	Male	622	656.26	408195.00
	Female	628	595.03	373680.00
Self-efficacy for Writing	Male	623	644.89	401764.50
	Female	624	603.15	376363.50
Self-efficacy for Reading	Male	623	654.05	407471.50
	Female	625	595.05	371904.50
Self-efficacy for Listening	Male	622	645.87	401730.50
	Female	622	599.13	372659.50

The results from Mann-Whitney Test Statistics in Table 5.3 indicate that male students reported significantly higher self-efficacy ($Mdn = 76$) than their female counterparts ($Mdn = 72$) in Speaking, $U = 176,174$, $z = -2.99$, $p = .003$. Male students also reported significantly higher self-efficacy ($Mdn = 68$) than female ($Mdn = 66$) in Writing, $U = 181,363.50$, $z = -2.04$, $p = .04$. Similarly, male students had higher self-efficacy ($Mdn = 75$) than female ($Mdn = 72$) in Reading, $U = 176279.50$, $z = -2.89$, $p = .004$ and higher self-efficacy for male ($Mdn = 65$) than female ($Mdn = 62$) in Listening, $U = 178906.50$, $z = -2.29$, $p = .02$. The findings showing female students reporting lower self-efficacy was not unexpected in the context of Rwanda as a

traditionally male dominated society. One of the possible reasons for these differences may be cultural, where some female students may perceive themselves as less capable than their male counterparts even when it is may not be necessarily true (see Section 6.5.2.1)

Another student demographic factor that was tested was the student age. As was done for the analysis on gender, a non-parametric test was used to measure a continuous dependent variable (i.e. the students' self-efficacy) against one independent variable (i.e. student age) which had ordinal data. A non-parametric test was found appropriate given the fact that this analysis treated both ordinal data (age) and a non-normally distributed data (students' scores on self-efficacy). The Kruskal-Wallis test was therefore used to measure the significance of differences in the students' self-efficacy in the four English language skills based on the students' age group (*13 and below, 14-15 and 16 and above*).

Table 5.4: Kruskal-Wallis Mean Ranks of the students' self-efficacy by age group

	Age group	N	Mean Rank
Self-Efficacy for Speaking	13 years and below	69	797.61
	14 to 15 years	514	688.37
	16 and above	653	544.58
Self-Efficacy for Writing	13 years and below	69	800.75
	14 to 15 years	514	700.67
	16 and above	651	532.41
Self-Efficacy for Reading	13 years and below	69	763.96
	14 to 15 years	513	697.01
	16 and above	653	540.50
Self-Efficacy for Listening	13 years and below	69	799.58
	14 to 15 years	511	691.48
	16 and above	651	537.30

The results of Kruskal-Wallis test displayed in Table 5.4 highlight the mean ranks of the students' self-efficacy in each of the four skills. An inspection of the mean ranks for the age groups in Table 5.4 shows that the younger group (*13 years and below*) had the highest self-efficacy levels across all the four language skills (see ranks in bold, in Table 5.4), with the older group (*16 years and above*) reporting the lowest self-efficacy in all the four language

skills. For the statistical significance of these differences, the Kruskal-Wallis test statistics confirmed that there was a statistically significant difference in the students' self-efficacy based on the three age groups for speaking skill, $\chi^2(2) = 65.07, p = 0.000$; for Writing skill, $\chi^2(2) = 83.35, p = 0.000$; for Reading skill, $\chi^2(2) = 67.57, p = 0.000$ and for Listening, $\chi^2(2) = 85.34, p = 0.000$. These results and the mean comparison in Table 5.4 seem to suggest that age was an important factor given the clear differences illustrated by the mean ranks. It is apparent from these findings that the older students felt the least self-efficacious of the three age groups. This may be a result of many factors including that some students in the older group were repeating the year or had enrolled in F2 only to be able to resit the required national examination before they could choose a major in upper secondary schools. For those students, having had unsatisfactory achievement before might have affected their ability beliefs thus rating their self-efficacy as low (see Section 6.5.2.2).

Table 5.5: Mann-Whitney Mean Ranks for students' self-efficacy by school learning mode

	Learning mode	N	Mean Rank	Sum of Ranks
Self-Efficacy for Speaking	Boarding	577	793.87	458063.00
	Day	674	482.29	325063.00
Self-Efficacy for Writing	Boarding	577	811.78	468396.50
	Day	671	463.46	310979.50
Self-Efficacy for Reading	Boarding	577	809.31	466972.50
	Day	672	466.74	313652.50
Self-Efficacy for Listening	Boarding	576	815.66	469821.50
	Day	669	457.12	305813.50

Differences were also observed for self-efficacy as a function of school learning mode (See Table 5.5). The surveyed schools were classified in two categories depending on the system of the school. Boarding school had students who lived in the school during the whole term time while students in day schools stayed in the school during the day and returned home after class. A variable was thus created on the basis of this classification to examine if a differences in the students' self-efficacy existed based on the two factors.

The results in Table 5.5 show that students in boarding school reported the level of self-efficacy nearly twice as high as that of students in day schools. The results of the Mann-

Whitney U test statistics revealed that the students in boarding schools reported significantly higher self-efficacy ($Mdn = 84$) than students in day schools ($Mdn = 66$) in Speaking, $U = 97,588$, $z = -15.21$, $p = .001$, and in the Writing skill where students in boarding schools showed higher self-efficacy ($Mdn = 78$) than those in day schools ($Mdn = 57$), $U = 85,523$, $z = -17.02$, $p = .001$. Similarly, the students in boarding schools reported higher self-efficacy ($Mdn = 84$) than students in day school ($Mdn = 63$) in Reading, $U = 87, 524$, $z = -16.73$, $p = .001$ and higher self-efficacy in Listening for boarding school students ($Mdn = 76$) than day school students ($Mdn = 51$), $U = 81, 698$, $z = -17.54$, $p = .001$. These results were not surprising given the differing contextual characteristics of the two systems. It is important to recall that the boarding schools were in most cases better resourced than day schools and usually received the highly performing students in the P6 national exams (see Section 6.5.3). Also, similar to the findings on self-efficacy as a function of age, these results indicate that the students reported higher self-efficacy levels in speaking and reading than in writing and listening.

Table 5.6: Mann-Whitney Mean Ranks of the students' self-efficacy by location of school

	Location of school	N	Mean Rank	Sum of Ranks
Self-Efficacy for Speaking	Rural	516	658.62	339848.00
	Urban	735	603.10	443278.00
Self-Efficacy for Writing	Rural	516	641.22	330869.50
	Urban	732	612.71	448506.50
Self-Efficacy for Reading	Rural	516	642.40	331477.50
	Urban	733	612.75	449147.50
Self-Efficacy for Listening	Rural	515	640.95	330089.00
	Urban	730	610.34	445546.00

Statistical tests were also used to examine whether differences in the students' self-efficacy existed based on the school location. In this study, the collected data was also arranged by school location. Data was gathered from schools both in urban and rural areas and information that was obtained was recorded under two separate variables of *Urban School* and *Rural School*. A school was considered rural when it was located in a rural administrative sector of a district or at 10 kilometres from the city centre. Schools located in administrative sectors officially labelled as urban or schools located in less than 10 kilometres from city centres were classified as urban. This school classification was done to verify if the students'

self-efficacy differed as a function of the location of the school. As can be seen in Table 5.6, students in Rural schools reported higher self-efficacy than students in Urban schools in all the four language skills. However, the Mann-Whitney test statistics was used to identify the skill in which the students' self-efficacy differed significantly. The results indicated that the difference in the students' self-efficacy was significantly higher only in the Speaking skill where the students in schools located in rural areas reported higher self-efficacy ($Mdn = 76$) than students in rural schools ($Mdn = 73$), $U = 172798$, $z = -2.68$, $p = .007$. No statistically significant difference was found between the students in Rural schools ($Mdn = 68$) and students in urban schools ($Mdn = 66$) for the Writing skill, $U = 180228$, $z = -1.37$, $n.s.$ The difference was also non significant in the Reading skill between Rural schools students ($Mdn = 74$) and Urban schools students ($Mdn = 72$), $U = 180136$, $z = -1.43$, $n.s.$ No significant differences was also found in the Listening skill between Rural schools students ($Mdn = 64$) and Urban schools students ($Mdn = 62$), $U = 180228$, $z = -1.48$, $n.s.$

Differences based on the school location were less expected because schools in rural and urban areas have many characteristics in common. The ministry of education in Rwanda encouraged educational and social cohesion by placing primary school leaving students to different secondary schools across the country. As a result, both rural and urban schools were perceived to represent a microcosm of the Rwandan society by enrolling students from different backgrounds and from all parts of the country. However, while the surveyed schools in rural areas included both day and boarding schools, it may be important to recall that the surveyed schools in Urban areas were all day schools (see Section 3.5.1). This may have had a central effect given the great significance of the school learning mode as a determining factor of the students' self-efficacy (see Table 5.5).

Analysis was also extended to the students' family background, looking at the level of education of their parents and examining whether it had any effect on the students' self-efficacy. The Kruskal-Wallis test was used to calculate the degree of significance of differences in the students' self-efficacy in the four language skills. Table 5.7 provides a breakdown of the mean ranks showing that there were apparent differences in the students' self-efficacy based on their parents' educational level. The results seem to suggest that the students' self-efficacy levels rose as a function of parents' education.

Table 5.7: Kruskal-Wallis Mean Ranks of the students' self-efficacy by parents' education

	Parents education	N	Mean Rank
Self-Efficacy for Speaking	Primary education	195	455.30
	Secondary education	463	545.25
	University education	585	738.31
Self-Efficacy for Writing	Primary education	192	452.78
	Secondary education	464	522.79
	University education	584	753.28
Self-Efficacy for Reading	Primary education	193	440.58
	Secondary education	464	526.02
	University education	584	756.09
Self-Efficacy for Listening	Primary education	192	436.48
	Secondary education	463	525.91
	University education	582	753.27

The K-S Test Statistics show that there was significant differences in the students' self-efficacy based on parents' education in Speaking: $\chi^2(2) = 124.64, p = 0.000$, in Writing: $\chi^2(2) = 156.97, p = 0.000$, in Reading: $\chi^2(2) = 164.50, p = 0.000$ and in Listening: $\chi^2(2) = 65.07, p = 0.000$. The mean rank displayed in Table 5.7 clearly indicates that the students whose parents were educated at the university level had the highest self-efficacy for all the four language skills. It also shows that the students whose parents' education was limited at the primary level had the lowest self-efficacy levels across all the skills. These were significant findings as they revealed the extent to which factors related to the parent's level of education can exert influence on the students' self-efficacy

It was also found that there was a statistically significant difference in the students' self-efficacy based on the status of the school. The self-efficacy levels of the students were compared according to whether the students studied in private, public or government aided schools. As demonstrated by the K-S mean ranks in Table 5.8, the students in private schools reported the highest levels of self-efficacy for all the four skills while the public school students reported the lowest.

Table 5.8: Kruskal-Wallis Mean Ranks of the students' self-efficacy by school status

	School Status	N	Mean Rank
Self-Efficacy for Speaking	Private	286	687.26
	Government aided	512	685.55
	Public	453	520.01
Self-Efficacy for Writing	Private	286	724.17
	Government aided	511	675.47
	Public	451	503.54
Self-Efficacy for Reading	Private	285	731.81
	Government aided	512	667.15
	Public	452	509.91
Self-Efficacy for Listening	Private	285	746.65
	Government aided	510	666.46
	Public	450	495.43

When the K-S tests were calculated to determine the significance of the differences illustrated in Table 5.8, a statistically significant difference was found between the three school status groups in Speaking: $\chi^2(2) = 61.12, p = 0.000$, in Writing: $\chi^2(2) = 82.90, p = 0.000$, in Reading: $\chi^2(2) = 78.00, p = 0.000$ and in Listening: $\chi^2(2) = 97.82, p = 0.000$. The mean rank in Table 5.8 indicates that students in private schools reported the highest levels of self-efficacy in all the four language skills while students in public schools reported the lowest self-efficacy. As was expected, the comparison of these three school categories on different factors revealed some key contextual differences that may have contributed to the observed high levels of students' self-efficacy in private school category. For example, while it was found that the level of parents' education had significant effect on students, it was found that up to 66% of students in Private schools had parents whose education was at the university level (see Section 3.5.1)

Table 5.9: Mann-Whitney Mean Ranks of Students' self-efficacy by teacher's gender

	Teacher's gender	N	Mean Rank	Sum of Ranks
Self-Efficacy for Speaking	Male	45	38.07	1713.00
	Female	21	23.71	498.00
Self-Efficacy for Writing	Male	45	38.04	1712.00
	Female	21	23.76	499.00
Self-Efficacy for Reading	Male	45	38.49	1732.00
	Female	21	22.81	479.00
Self-Efficacy for Listening	Male	45	37.89	1705.00
	Female	21	24.10	506.00

Data on the students' self-efficacy was also analysed to examine if and the extent to which the gender of the teacher was related to the self-efficacy beliefs of the students. As also shown by the mean ranks in Table 5.9, the students taught by male teachers reported higher self-efficacy across all four language skills. The Mann-Whitney U test results indicated that the self-efficacy levels reported by students taught by male teachers was significantly higher ($Mdn = 73$) than female teacher taught students ($Mdn = 61$) in Speaking: $U = 267, z = -2.82, p = .005$. Students taught by male teachers also reported significantly higher self-efficacy in Writing ($Mdn = 68$) than those taught by female students ($Mdn = 56$), $U = 268, z = -2.81, p = .005$. The students' self-efficacy levels in Reading and Listening were also found significantly higher ($Mdn = 73$ and $Mdn = 65$ respectively) among male teacher taught students than female teacher taught students ($Mdn = 61$) for reading: $U = 248, z = -3.10, p = .002$ and ($Mdn = 51$) for Listening: $U = 275, z = -2.71, p = .007$.

Overall, the results from the Mann-Whitney and Kruskal-Wallis tests indicated that the students' self-efficacy levels significantly differed as a function of many factors including mainly the demographic related factors (age, gender) and the school related factors such as the school status, the school location or school learning mode. Some of these factors seemed to be more important than others. In particular, the results from the statistical significance tests in tables 5.4 and 5.5 suggest that age and school learning mode may be the most important factors. In the following sections, more facts from the students' interviews on their beliefs about their self-efficacy are presented and some interview excerpts are used for higher precision. The letter "R" is used to refer to the 'researcher' and letter 'S' to refer to 'student'.

The results from four students' focus group interview with the students seem to be consistent with the results from the questionnaire survey on the students' self-efficacy. The students were asked to say the number that they thought corresponded to the extent to which they believed that they could do specific tasks effectively in speaking, writing, listening and reading in English. As in the questionnaire survey (See appendix 1a, part III), the students were given examples of tasks and were asked to indicate their self-efficacy level by choosing a number from 0 to 100, to show the extent to which they could complete such a task accurately in English (See appendix 2a). Their answers seem to echo the results from the questionnaire survey by reemphasizing their highest self-efficacy in Writing and Reading (see Table 5.10). In

addition, the majority of the interviewed students also indicated that they felt least self-efficacious in listening while the ratings on speaking remained around 50% as indicated in the sample interview excerpt below from Focus Group A.

Researcher: From 0 to 100, to what extent can you speak good English explaining the importance of literacy... to what level? ...

S1: at 60%

S2: For me it's 50%

S3: Me too it's 50%

Researcher: Ok 50% means it's in the middle (hand gestures to show balance)

S4: For me it's 60%

S5: It's 40%

S6: 50%

S7: 50%

S8: 60%

The overall inspection of the students' self ratings shown in Table 5.10 indicates that the respondents in Focus group A felt more self-efficacious in writing ($\bar{x}=65$) and reading ($\bar{x}=63$). The results from two other focus groups show the same tendency, that students' self-efficacy was higher in reading and writing. These results match some of the findings from the questionnaire survey (see Table 5.7) where students reported higher self-efficacy in writing and reading. However, as it was also reported in some other data sets (see Table 5.9 and Table 5.6), the focus group students reported the lowest self-efficacy in Listening.

Table 5.10: Students self-efficacy (sample from focus group A)

Student	SE for Speaking	SE for writing	SE for Listening	SE for Reading
	%	%	%	%
S1	60	70	65	90
S2	50	60	50	60
S3	50	55	45	60
S4	60	75	40	45
S5	40	60	55	60
S6	50	75	60	60
S7	50	70	40	60
S8	60	60	60	65
Mean	53	65	52	63
Mode	50	60	40,60	60

One of the main aims of the students' interview was to gather more information to compare with data from the questionnaire survey. Therefore, follow up questions were asked to gather more details about the students reported self-efficacy in order to identify the reasons upon which students based their self-efficacy ratings (See appendix 10). In their responses, the students referred to classroom assessment emphasizing that it was easy for them to predict their performance. However, students clarified that their confidence level was particularly high in tasks that were similar to what they did in their classrooms. The following interview extract from student Focus Group A provides some insights into what the students' own judgement of their self-efficacy was reportedly based on.

Researcher: Ok, ah why? Let's everyone remember the number you've just given, for example you said 60% right? Why do you think you can do it at 60%?

S1: ...because I feel that there are some things that I cannot say well but I think there is a lot I can say well.

Researcher: Ok another one to tell us why you chose that number?

S7: ...because I am aware of my English, I feel I can do it at 50%. Because it may be wrong to say 60 or 70% while I actually can't speak good English.

S6: The reason why I said 55% is because when you look at the way we study English here, we focus on written English. So speaking becomes a challenge to us... but we can write.

S5: I said 40%, we often meet challenges for example you leave primary with limited knowledge in some areas and only start to understand in secondary school.

The statement by S6 in the excerpt above appears to suggest that he measured his ability to successfully complete tasks in speaking at 55% level because their learning was more focused on written English than speaking. This was also repeated in other groups where the students' comments gave the impression that writing and reading tasks were often done during classroom assessment which increased the students' confidence that they could perform tasks in these skills more easily than speaking and listening. In addition to what happened in the classroom, however, Student S5 gave an example of other factors that affected the students' self-efficacy beliefs, citing the inability to speak English due to poor previous experience in primary school. This was not surprising because many of the teachers in primary schools in Rwanda had little or no experience and training in teaching in English (Pearson, 2013). It also highlights some students' perceived weakness which may have affected the way they responded to the self-efficacy questions in the survey.

5.2 The relationship between the students' self-efficacy and the students' perceptions

Data on the students' perceptions of the teachers' classroom assessment practices was analysed to help answer sub-research question 2.2 regarding the relationship between the students' perceptions and their self-efficacy: *How is the students' self-efficacy for the four English language skills related to the students' perceptions of the teachers' assessment practices?* The Spearman correlation tests were used to measure the extent to which each of the 8 independent variables i.e *Congruence, Authenticity, Student consultation, Transparency, Diversity, Feedback, Fair marking and Frequency* was related to the four dependent variables i.e. the students' self-efficacy in Speaking, Writing, Reading and speaking.

Table 5.11: Spearman's rho correlation between students' self-efficacy (SE) and students' perceptions

			SE for Speaking	SE for Writing	SE for Reading	SE for Listening
Spearman's rho	Congruence with learning material	Correlation Coefficient	.162**	.206**	.190**	.190**
		Sig. (2-tailed)	.000	.000	.000	.000
		N	1239	1236	1237	1233
	Task authenticity	Correlation Coefficient	.250**	.266**	.266**	.261**
		Sig. (2-tailed)	.000	.000	.000	.000
		N	1221	1218	1219	1215
	Transparency of assessment aims	Correlation Coefficient	.357**	.361**	.353**	.347**
		Sig. (2-tailed)	.000	.000	.000	.000
		N	1240	1238	1239	1235
	Teacher feedback	Correlation Coefficient	.142**	.145**	.129**	.112**
		Sig. (2-tailed)	.000	.000	.000	.000
		N	1237	1235	1236	1232
	Teacher fair marking	Correlation Coefficient	.091**	.110**	.115**	.102**
		Sig. (2-tailed)	.001	.000	.000	.000
		N	1247	1245	1246	1242
	Frequency of assessment	Correlation Coefficient	.199**	.208**	.198**	.190**
		Sig. (2-tailed)	.000	.000	.000	.000
		N	1246	1244	1245	1241

Correlation is significant at the 0.01 level (2-tailed).**

Correlation is significant at the 0.05 level (2-tailed).*

The Spearman correlation coefficient is a non-parametric test that can be used when data violates the parametric tests assumptions as in case of non-normally distributed data. It is an equivalent of the parametric Pearson correlation coefficient and is used to measure the level of association between a continuous dependent variable and an ordinal data (Field, 2009). In this analysis the independent ordinal variables were the students' perceptions and the dependent continuous variables were the students' self-efficacy in each language skill. The correlational attributes between these explanatory variables (perceptions) and the explained variable (students' self-efficacy) was analysed and the results indicated that six of the eight perception variables were significantly correlated with the students' self-efficacy in all four skills (see Table 5.11). On the other hand, the results of the Spearman rho test indicated that the students' perceptions on *student consultation* and *task diversity* did not correlate with self-efficacy in any of the English language skills.

Although statistically significant at $p < .01$, the low correlational coefficients presented in Table 5.11 indicate that a weak relationship existed between the perception variables and the students' self-efficacy for both receptive and productive skills. These results show that *congruence* between assessment tasks and learning material was related to the students' self-efficacy in all skills. In particular, the students' self-efficacy in the Writing skill appears to be the most highly correlating with congruence ($r=.20, p = .000$) where the positive correlation coefficient suggests that the more students believed that the assessment tasks agreed with the learning material, the more they felt self-efficacious. Specifically, the calculated coefficient of determination (R^2*100) indicate that congruence helped explain 4% of the students' self-efficacy in writing. Significant correlation was also found between *Task Authenticity* and the students' self-efficacy in all the four skills. The results in Table 5.11 show a correlation coefficient of $r=.26$ between the students' score on authenticity and their self-efficacy in Writing, Reading and Listening. As for *congruence*, this coefficient indicates a small but meaningful correlation, suggesting that students who strongly agreed that classroom assessment tasks were authentic also reported higher self-efficacy in Writing, Reading and Listening skills. The Spearman rho coefficient of .26 also meant that 6% of the students' self-efficacy in Writing and Reading could be explained by *Authenticity* of assessment tasks.

One of the important findings of the Spearman correlation test presented in Table 5.11 is the strength of the relationship between the students' perceptions on *transparency* of assessment objectives and the students' self-efficacy. *Transparency* was the only variable that had the value of Spearman correlation (r) higher than .29. The correlation coefficients above .29 indicate 'medium' relationship strength (Cohen, as cited in Pallant, 2010). Also, coefficients in a range between $r = .35$ and $r = .65$ are classified as statistically significant for possible prediction of the dependent variable (Borg, 1963; Cohen et al. 2007). As illustrated in Table 5.11, the r coefficient for perceptions on *Transparency* was above $r = .35$ for Speaking, Writing and Reading, meaning that up to 12% of variance in the students' self-efficacy in each of these three skills can be explained by the students' perceptions of transparency of assessment objectives. In other words, these results seem to imply that students who strongly believed that the assessment objectives of their classroom tasks were clear also demonstrated higher levels of self-efficacy in all the skills including listening ($r = .34, p = .000$). The relatively higher relationship between *Transparency* and self-efficacy was expected given the findings from other studies (e.g Alkharusi, 2013; Dorman et al., 2006). It seems realistic to predict that understanding of one's own goal helps in the evaluation of one's own achievement which in turn has an impact on one's beliefs about their ability to accomplish tasks.

Relationship was also found between the students' perceptions of the teachers' feedback and the students' self-efficacy. However, this was a 'weak' relationship as the size of the calculated Spearman correlation (r) fell under $r = .30$. The results in Table 5.11 indicate that the highest r value for Teacher feedback was $r = .14$ for the relationship with Self-efficacy in Writing and Speaking. These were to an extent unexpected given the important role that teacher feedback plays in recognising the learner's strengths as well as identifying their weaknesses and suggesting ways for improvement. One of the reasons of the low correlation may arguably be the nature of feedback provided to these learners and the learners' own limited awareness of the teachers' feedback alternatives (see Section 6.1.3). Lower but significant relationship was also found between the students' self-efficacy and the students' perceptions of *teacher fairness in marking*, and the students' perceptions of the *frequency of assessment*. Looking at the results in Table 5.11, it is evident that the students' perceptions on *teacher fairness in marking* had the lowest correlation coefficient for its relationship with Writing, Reading and Listening. Although significance level seems to suggest that there is correlation with Speaking, the (r)

coefficient of .09 indicates that there is no relationship between the two variables. In large sample data, very small Spearman correlations can be statistically significant and it may hence be necessary to put more emphasis on the size of the (r) coefficient and the shared variance between variables (Pallant, 2010).

In a summary, the results of the Spearman correlation tests indicated that some degree of relationship existed between the students' perceptions of the teachers' practices in the classroom and the students' self-efficacy in the four language skills. The sum of the coefficients of determination for each perception variable indicates that the students' perceptions of the classroom assessment can help explain up to 26% of the variance in the students' self-efficacy for the Speaking skill, 30% of the variance in the students' self-efficacy for Writing, 29% of the variance in the students' self-efficacy for Reading and 27% of the variance in the students' self-efficacy for Listening. The results also suggest that transparency had the highest coefficient indicating its strongest relationship with the students' self-efficacy in all the four English language skills.

5.3 Relationship between the students' self-efficacy and the teachers' assessment practices

The results presented in chapter four offered some important information of the classroom assessment practices of the teachers of English at the lower level of secondary schools in Rwanda. Specifically, the results provide important insights into the teachers' practices in relation to the forms of assessment and the methods that they used to provide feedback to their students. In this section, I revisit these results and examine whether and the extent to which the use of assessment forms by teachers was related to the students' self-efficacy. In addition to the analysis of the relationship between the students' self-efficacy and the students' perceptions of teachers' classroom assessment presented in the previous section, this section seeks to find answers to research question two precisely on its sub-question three: *How is the students' self-efficacy for the four English language skills related to the teachers' assessment practices?* In order to answer this question, the Spearman correlation analysis was used to measure the nature and strength of relationship between specific forms of assessment and the students' self-efficacy in each of the four English language skills. In addition, the teachers' methods of

providing feedback were also correlated with the students' self-efficacy to measure the level of association between the two variables. This was done to examine whether the frequent use of a particular method correlated with the level of the reported student self-efficacy in English Speaking, Writing, Reading and Listening skills. As data on the forms of assessment used in the classroom was obtained from teachers (self-report) and their students, both data sets were analysed and compared in order to identify a common trend between both results.

Based on Bandura's (1997) Social Cognitive Theory –SCT, specifically on mastery and vicarious experiences as sources of self-efficacy component, this question was formulated to examine if the students' self-efficacy beliefs varied as a function of their linguistic performance experiences during classroom assessment. Assessment practices happening in the classroom, considered here as a micro social setting in the context of SCT, were examined and measured against the students' beliefs for significance of relationship. According to Bandura (1997), of the four sources of self-efficacy, experience of failure or success (mastery experience) and witnessing others succeed or fail (vicarious experience) are the two major sources that increase or decrease people's self-efficacy (See section 2.5.2). In the academic field, attainable learning goals that enable students to experience achievement through successful completion of assessment tasks, supported with teachers positive feedback (praise and encouragement) are expected to boost students' self-efficacy or to erode it in the case of the opposite (Schunk and Pajares, 2002).

Based on SCT and in the context of this study, it would be expected for example that the students who were regularly and successfully engaged in *performance* assessment tasks such as debates and oral presentations would demonstrate higher levels of self-efficacy for speaking than students whose learning and assessment was focused on grammar. In other words, the study examined whether students taught by teachers who mostly used *performance* assessment felt more self-efficacious than students who were taught by teachers who often used *paper-and-pencil* forms of assessment (See section 4.1.1). Thus, the Spearman correlation tests were carried out to examine whether students' level of self-efficacy was related to the use of a given type of assessment forms or the method of providing feedback to students. Data was also triangulated with qualitative information obtained from classroom observation and interviews.

5.3.1 Relationship between the use of assessment forms and the students' self-efficacy

The Spearman correlation test was used to examine the relationship between the teachers' use of *forms of assessment* as independent variables on the one hand, and the *students' self-efficacy* as dependent variables on the other hand. The teachers' median score on each sub-scale of assessment forms pertaining to each language skill (See Table 4.1) was calculated and correlated with the mean score of the students' self-efficacy scores in the corresponding skill in each class. The test of the internal consistency of the sub-scales of assessment forms demonstrated acceptable levels of reliability with Cronbach's Alfa ranging from .65 for assessment forms related to speaking, to .78 for assessment forms related to writing. For example, to measure the relationship between the teachers' use of performance-based assessment forms and the students' self-efficacy in speaking, the median of the score on how frequently the teachers used the suggested assessment forms related to speaking was correlated with the aggregated classroom score of the students' self efficacy in speaking. Each median score from each sub-scale was correlated with the students' self-efficacy for the corresponding language skill. In addition, as data on the forms of assessment used in the classroom was also collected from the students (see Table 4.2), the analysis also examined the students reported data for comparison with the results from the teachers' self-report. The correlational analysis indicated that a weak relationship existed between the use of some performance-based forms of assessment and the students' self-efficacy.

5.12: Correlation between the teachers' use of assessment forms related to speaking and the students' self-efficacy in Speaking (Data from teachers)

			Use of performance assessment related to speaking	Student self-efficacy in Speaking
Spearman's rho	Use of performance assessment related to speaking	Correlation Coefficient	1.000	.132
		Sig. (2-tailed)	.	.334
		N	56	56
	Student self-efficacy in Speaking	Correlation Coefficient	.132	1.000
		Sig. (2-tailed)	.334	.
		N	56	56

When teachers' self-reported assessment forms related to speaking were correlated with the students' self-efficacy in speaking, the results from the Spearman rho correlation test showed that no relationship existed between these two variables (See Table 5.12). As stated earlier in this section, some minimum of relationship between the two variables was expected on assumptions that giving students the opportunities to practice speaking would have influence on their self-efficacy in speaking.

5.13: Correlation between the teachers' use of assessment forms related to speaking and the students' self-efficacy in Speaking (Data from students)

		Use of performance assessment related to speaking	Student self-efficacy in Speaking
Spearman's rho	Use of performance assessment related to speaking	Correlation Coefficient	1.000
		Sig. (2-tailed)	.
		N	1247
	Student self-efficacy in Speaking	Correlation Coefficient	.241**
		Sig. (2-tailed)	.000
		N	1247

** . Correlation is significant at the 0.01 level (2-tailed).

These results were therefore unpredicted and raised more questions including questions on the reliability of the teachers' responses in the questionnaire (see Section 6.5.4 and Section 7.3). On the other hand, the analysis of data from the students indicated that the frequent use of oral presentations in the classroom was positively correlated with the students' self-efficacy. As shown in Table 5.13, the output of the Spearman rho test demonstrates a statistically significant correlation between these two variables ($r = .24$, $p = .000$). Nonetheless, this low correlation coefficient ($r = .24$) meant that there was a weak shared variance of $R^2 = .05$ ($.24 \times .24$), suggesting that only 5% of the students' self-efficacy for speaking could be explained by the teachers' use of performance assessment forms related to speaking in the classroom.

The mismatch between the teachers' reported information on the one hand and the students' report on the other hand, was noted in some data sets analysis. However, this may not be surprising as various factors may have influenced the answers of some respondents on a

number of questions (see section 7.3). Although the teacher' survey seemed to show a balanced use of assessment forms between performance based and paper and pencil based assessment, triangulation with other sets of data appeared to suggest a rather much dominant use of paper and pencil based assessment (see section 6.1.1). This, in addition to other assessment related factors might have led to the current divergent results.

5.14: Correlations between the teachers' use of assessment forms related to writing and the students' self-efficacy in Writing (Data from teachers)

		Student self-efficacy in Writing	Use of performance assessment related to Writing
Spearman's rho	Student self-efficacy in Writing	Correlation Coefficient	1.000
		Sig. (2-tailed)	.
		N	61
	Use of performance assessment related to Writing	Correlation Coefficient	.338**
		Sig. (2-tailed)	.008
		N	61

** . Correlation is significant at the 0.01 level (2-tailed).

The Spearman correlation test was also used to measure the level of relationship between the teachers' use of assessment forms related to writing and the students' self-efficacy in the writing skill. As was done for speaking, the teachers' median score on the subscale of assessment forms related to writing was correlated with the average mean score of the students' self-efficacy in the writing skill in each class. The results from both teachers' survey (Table 5.14) and students' surveys (Table 5.15) indicated that a statistically significant relationship existed between the two variables. Results in Table 5.14 indicate that teachers who reported frequent use of performance based forms to assess writing had students with higher self-efficacy in writing ($r = .34, p = .008$). The calculated coefficient of determination ($R^2 = .11$) suggest that 11% of the differences in the students' self-efficacy in writing can be explained by the use of performance based assessment forms related to writing. A statistically significant correlation between these two variables was also obtained ($r = .18, p = .000$) when the students' score on the frequency of use of assessment forms related to writing was correlated with their self-efficacy in writing (see Table 5.15). This shows that, according to the students' responses, only 3% of the variance was shared by the two variables. The Spearman correlation coefficients

shown in Tables 5.13 and 5.14 demonstrate a slight relationship as they are below $r=.35$ and cannot be used for predictions (Cohen et al., 2007). However, the statistical significance levels obtained from the analysis of the two sets of data (teachers and students data) highlight the existence of a relationship between the use of assessment forms related to writing and the students' self-efficacy in the Writing skill.

5.15: Correlations between the teachers' use of assessment forms related to writing and the students' self-efficacy in Writing (Data from students)

			Students' self-efficacy in Writing	Use of performance assessment related to Writing
Spearman's rho	Students' self-efficacy in Writing	Correlation Coefficient	1.000	.182**
		Sig. (2-tailed)	.	.000
		N	1236	1236
	Use of performance assessment related to Writing	Correlation Coefficient	.182**	1.000
		Sig. (2-tailed)	.000	.
		N	1236	1236

** . Correlation is significant at the 0.01 level (2-tailed).

Unlike in speaking and writing, no statistically significant correlation was found between the teachers' use of assessment forms related to reading and the students' self-efficacy in the reading skill (See the results of correlation test in appendix 13). The results from the analysis of data from teachers ($r=-.05$, $p=.69$) and the results from the analysis of data from the students ($r=-.07$, $p=.01$) both indicated that no relationship existed between these two variables. These results were also unexpected given the reported high levels of students' self-efficacy in reading and the reported high frequency of use of assessment forms related to reading. One possible explanations of this incongruence may be that other assessment related factors other than the assessment forms per se may be more important and more influential to the students' self-efficacy. For example, the students' achievement score in reading was found to be associated with the students' self-efficacy (Smith et al., 2012). As a result, students may judge their ability based only on their level of achievement measured by scores on reading assessment tasks and not on their practical experience during assessment related to reading. Similarly, the Spearman Rho test results on the correlation between the students' self-efficacy for listening and the teachers' use of the forms of assessment related to listening did not show any

statistically significant relationship (See appendix 14). The results from the analysis of data from both the teachers ($r=.31$, $p=.06$) and from the students ($r=.02$, $p=.43$) show low correlation coefficients and non statistically significant ‘ p ’ value. This was highly predictable given that listening was constantly rated by both the teachers and the students as the least assessed skill.

5.3.2 The relationship between the use of feedback methods and the students’ self-efficacy

Further analysis was done to examine the extent to which the students’ self-efficacy was related to the methods through which they received feedback about their performance on assessment tasks in the classroom.

5.16: Correlation test between the teachers’ use of feedback methods and the students’ self-efficacy

			SE for Speaking	SE for Writing	SE for Reading	SE for Listening
Spearman's rho	Student peer feedback	Correlation Coefficient	.214	.208	.161	.150
		Sig. (2-tailed)	.104	.114	.223	.257
		N	59	59	59	59
	Collective verbal teacher feedback	Correlation Coefficient	-.112	-.100	-.095	-.146
		Sig. (2-tailed)	.389	.443	.465	.263
		N	61	61	61	61
	Individualized verbal teacher feedback	Correlation Coefficient	-.192	-.141	-.141	-.135
		Sig. (2-tailed)	.136	.273	.273	.295
		N	62	62	62	62
	Written collective teacher feedback (on the chalkboard)	Correlation Coefficient	-.188	-.123	-.118	-.080
		Sig. (2-tailed)	.148	.344	.364	.539
		N	61	61	61	61
	Individualized written teacher feedback	Correlation Coefficient	-.073	-.038	-.035	-.040
		Sig. (2-tailed)	.567	.770	.785	.755
		N	63	63	63	63
	Test results only (marks)	Correlation Coefficient	.040	.045	-.007	.030
		Sig. (2-tailed)	.755	.725	.956	.815
		N	63	63	63	63

Investigation on the methods used in the classroom sought to gather data on how frequently the students received information about their performance through verbal teacher comments, written teacher comments, collective or individualised teacher comments, through peer comments or through marks. Each method was correlated with the students' self-efficacy in each language skill. The Spearman correlation test found that there was no correlation between the teachers' reported use of feedback methods with the students' self-efficacy in any of the four language skills (see Table 5.16). As indicated in Table 5.16, analysis of data from the teachers' self-report on the methods of providing feedback used in the classroom failed to establish a relationship between the reported methods and the students' self-efficacy. This was less expected given some study findings that have identified predictive relationship between the teacher's feedback and the students' self-efficacy (e.g. Mehregan, 2014; Naderi, 2014; Thomas et al., 2012). In addition, while the analysis of data from teachers do not show any relationship between these two variables, the analysis of data from the students (see Table 5.11) seems to show a trend in the students' self-efficacy differences based on their perceptions of the teachers' feedback.

5.17: Differences in the students' self-efficacy based on their perceptions of teacher feedback

	I always get enough feedback from my teacher after assessment tasks.	N	Mean Rank
Self-efficacy for Speaking	Strongly disagree	63	552.19
	Disagree	211	575.39
	Agree	576	588.37
	Strongly agree	387	699.24
Self-efficacy for Writing	Strongly disagree	63	519.45
	Disagree	211	582.98
	Agree	575	587.43
	Strongly agree	386	698.77
Self-efficacy for Reading	Strongly disagree	63	544.77
	Disagree	211	580.48
	Agree	576	592.97
	Strongly agree	386	689.42
Self-efficacy for Listening	Strongly disagree	63	547.20
	Disagree	211	599.40
	Agree	574	585.06
	Strongly agree	384	684.27

As shown in Table 5.17, The Mean Ranks from the Kruskal-Wallis test of differences in the students' self-efficacy indicates that the students who strongly agreed that they received sufficient feedback from their teachers reported the highest self-efficacy in all the four English language skills. The differences between groups were found statistically significant in speaking ($\chi^2 (3) = 29.1, p = 0.000$) in writing ($\chi^2 (3) = 30.8, p = 0.000$), in reading ($\chi^2 (3) = 23.3, p = 0.000$) and in Listening ($\chi^2 (2) = 21.2, p = 0.000$). It can be seen in Table 5.17 that clear differences existed between groups in all skills except for listening where major differences seem to be between students who reportedly received no feedback and those who reported receiving enough feedback from teachers.

Although data from the students contained limited information in terms of the nature of feedback provided by the teachers, the results from its analysis demonstrated yet another evidence of the divergence between the students' and the teachers' responses. These divergences were also noted in the findings from the analysis of assessment forms used in the classroom (see Section 4.1.1). As explained earlier, this may have been a result of 'social desirability response bias' (see section 3.3.1) on the side of teachers who may have provided information that would convey favourable image of themselves to the researcher (Marsden and Wright, 2010). Nonetheless, the use of different instruments to collect data from different respondents allowed for the verification and validation of information through comparison of data from different sources.

5.4 Summary of the chapter

Chapter five reported on the students' perceptions of their teachers' classroom assessment practices and the students' self-efficacy for reading, listening, speaking and writing English. The chapter also presents results on the relationship between the students' self-efficacy in the four English language skills on the one hand and the teachers' use of assessment forms namely the use of performance assessment and the use of paper and pencil based forms of assessment on the other hand. The results indicate that the surveyed students reported positive perceptions of the teachers' assessment practices. The students also reported higher self-efficacy in all the

four language skills than was predicted given the consensus literature that students demonstrated little confidence in communicating in English. The results in chapter five also show that a certain degree of relationship exists between the use of assessment forms for writing and speaking and the students' self-efficacy in the corresponding skills. It is important to recall that inconsistency was observed in the results from data from students and data from the teachers. This is thought to have emerged from some issues such as the social desirability bias on the side of the teachers who might have wanted to be positively judged by the researcher. However, the collection of different sets of data using different instruments allowed for data triangulation which led to the overall conclusion that a marginally significant relationship existed between the use of some classroom assessment forms and the students' self-efficacy. Nonetheless, the results point to the important significance of the influence of other factors including the learners' age and gender and school based factors.

Chapter 6: Discussion

6.0 Introduction

This study set out to explore the classroom assessment practices of the teachers of English at the lower level of secondary schools in Rwanda. The focus on the teachers' assessment practices was particularly put on the forms of assessment and the methods of providing feedback to students that teachers used in their everyday classroom practices. Other assessment practices were also investigated namely the purpose and timing of teachers' assessment, the content focus of assessment activities as well as sources of assessment tasks used in the classroom. The study also aimed to examine the relationship between the teachers' assessment practices, especially the teachers' use of assessment forms and feedback methods, and the students' self-efficacy for using English in both productive skills (Writing and speaking) and receptive skills (Reading and listening). Chapter four has presented the results from the analysis of data that was collected to answer the first research question regarding the teachers' assessment practices and the students' perceptions of their teachers' assessment practices. In chapter five, the findings on the students' self-efficacy in the four English language skills were presented as well as the results on the relationship between the students' self-efficacy and the teachers' assessment practices described in chapter four.

In this chapter, the results from the analysis of data on each research question as presented in chapter four and five are discussed further by examining the findings in relation to the existing literature on classroom assessment practices and learners' self-efficacy in educational settings in general and in EFL context in particular. The chapter starts with a review of findings on research question one regarding the teachers' assessment practices. The discussion focuses on the results regarding the forms of assessment that teachers reportedly used in the classroom and the results on the teachers' methods of providing feedback to the students. The chapter also briefly highlights the importance of the findings on the students' perceptions of the teachers' assessment practices. Discussion from Section 6.4 focuses on the findings on research question two and develops a general understanding of the students' self-efficacy and its relationship with students' perceptions and with teachers' assessment practices.

6.1 Teachers' preferred assessment practices

In light of the current literature on the forms of assessment that teachers use for their classroom based assessment (see section 2.2.4) and based on my own personal experience as a former teacher in Rwandan secondary schools, it was hypothesized that the teachers' classroom assessment practices at the surveyed lower level of secondary schools in Rwanda were based more on *traditional* assessment than *performance* based assessment with the use of marks and grades as the teachers' main methods of providing feedback to students. It was also expected that, given the findings of the previous studies on the teachers' assessment practices (see Sections 2.3), the teachers assessed for summative purposes, focused their assessment on grammar and relied on external sources for their classroom assessment. In the following sections, the findings of the current study are discussed further and put in context in order to understand their significance vis-a-vis the current literature.

6.1.1 Prevalent use of *paper and pencil* based forms of assessment

Sub-research question 1.1 on the everyday classroom assessment used by teachers in the surveyed schools at the lower level of Rwandan secondary schools sought to explore the preferences for the teachers in terms of assessment forms used to assess students in the classroom. A particular distinction was made between *performance* based and *paper and pencil* based assessment (see section 2.2.4). As highlighted in chapter four, it was found that overall, teachers of English at the surveyed secondary schools preferred to use the *paper and pencil* based assessment more frequently than *performance assessment* (see section 4.1.1). It was revealed that forms such as '*gap fill*' and '*multiple choice*' were common and frequently used by most teachers (see Table 4.1 and Table 4.2). The current results seem to agree with the findings of other studies in which the controlled methods of assessment have been found to dominate the classroom assessment (Cheng et al., 2004; Frey and Schmitt, 2010; Tante, 2010; Tsagari and Pavlou, 2009). In their investigation of assessment methods and procedures of EFL teachers in public schools in Cyprus, Tsagari and Pavlou (2009) found that above 80% of the 191 surveyed teachers reported using paper and pencil based assessment for their classroom assessment. Teachers reported that they frequently used items such as true/false and multiple answer choice to assess their students. The same results are reflected in the current study as

shown in Table 4.1, where the top three highly rated forms of assessment that teachers reported as being frequently used in their classroom are all paper and pencil based.

Frequent use of paper and pencil forms was also confirmed by the results from the classroom observations where no performance assessment tasks were recorded in any of the seven observed classrooms (see Table 4.4). Teachers often used white chalk to write assessment tasks on the chalkboard and students used paper and pen to write down their short answers. Nonetheless, it is important to note that each of the actual classrooms was observed only once and for an average of 40 minutes for reasons mainly relating to the schools' timetables and the researcher's limited time (see Section 7.3 on study limitations). Therefore, it could arguably be said that this relatively limited classroom observation and the short length of the observed lessons did not allow enough opportunities for more assessment evidences to be captured. However, the comparison of results from the analysis of data from classroom observation was compared to the results from the analysis of other data sets seemed to confirm the frequent use of *paper and pencil* forms. For example, the teachers' answers on the specific period of time that they used each of these assessment forms indicated that they used *performance* assessment at the end of instruction more frequently than *paper and pencil* based assessment. This may indicate that the *performance* assessment was not frequently used for everyday classroom assessment but more used for summative purposes at the end of a specific instructional period (see Section 4.2.2). This was also stressed during interviews with teachers where some teachers indicated that performance assessment was used for few specific term periods due to time constraint.

Other important findings that seem to confirm the teachers' preference for *paper and pencil* assessment to *performance* assessment was the teachers' lack of interest in the use of rubrics. Data from teachers' interviews and classroom observation indicated that the majority of teachers did not use rubrics in their classroom based assessment. As presented in section 4.1.2, only 14% of the interviewed teachers said that they had used rubrics for their classroom assessment and 86% indicated that they were not familiar with rubrics and were unaware of the role that rubrics can play in assessment. In the context of this study, it appears that the rare use of rubrics may be related to the teachers' use of a limited range of *performance* assessment where criteria-based methods would need to be used as Arter and McTighe (2001) explain:

“Since the performance assessments generally do not yield a single correct answer or solution method, evaluations of the student products or performances are based on judgments guided by criteria.” (p. 180). These students’ products and performances may consist of an essay, a long term project, a research paper or another relatively complex assignment (Andrade, 2000). Therefore, as the teachers predominantly used paper and pencil forms for their classroom assessment that usually required single correct answers, they may have felt no need to use criteria based assessment where Rubrics would be needed. Instead, some teachers explained that they avoided using assessment based on subjective answers especially for peer assessment.

The prevalent use of *paper and pencil* based forms in the classroom-based assessment can be influenced by a number of factors. Teachers often prefer to use *paper and pencil* assessment for different reasons including that assessment based on these forms can be scored easily and objectively contrary to performance based assessment that can be subjective, hard to score and very time consuming (Clapham, 2000; Elliott et al. 2000, Yang, 2008). Unlike in performance based assessment, scoring a *true and false* test or *sentence completion* items that usually require one right answer can be straightforward and less time consuming than scoring an essay as emphasized by Elliot et al. (2000): “Good multiple choice items are difficult to prepare but can be scored easily and objectively. Essay tests, on the other hand, are relatively easy to prepare but extremely difficult to score.” (p. 218). The surveyed teachers in the current study did not point to the difficulty of scoring as a hindrance to their use of performance assessment. However, the responses of some teachers on why they focused on *paper and pencil* assessment seemed to suggest that these assessment forms were preferred because they could be marked with little difficulty. As was inferred from the comments made by some interviewed teachers, using paper and pencil forms was also more common for teachers who taught classes with a large number of students. The use of paper and pencil based assessment in large class size environment has been highlighted in some research studies. For example, in their investigation of factors influencing Chinese secondary school teachers of English, Chen and Sun (2015) found that many of the 350 surveyed teachers who taught large classes reported using paper and pencil forms because they found it easy to mark the students work. Similarly, as some interviewed teachers explained, teachers who taught large classes in Rwandan

secondary schools may consider that scoring many students using performance assessment is challenging and therefore resolve to stick with the use of paper and pencil based assessment.

In fact, the perceived challenges that teachers associate with performance based assessment have been found to be very influential in the teachers' choice of classroom assessment forms. In her investigation of 425 EFL Taiwanese teachers from grade one to grade six on the factors that influenced their use of multiple classroom assessment practices, Yang (2008) found that the teachers' perceived difficulties such as time constraints, difficulty with classroom management and the parents' doubts on the objectivity of grading negatively affected their use of alternative (performance based) assessment. Apart from the parents' doubt on the grading objectivity, all the challenges identified in Yang's study were also mentioned by participants in the current research (see section 4.1.1). The teachers may have not mentioned challenges related to the objectivity of their grading because they normally reported to the school administration, not directly to the parents.

Yang's study also found that teachers' assessment education and training had some effect on their use of classroom assessment practices. In the current study, the results showed that half of the surveyed teachers had had one or no in-service assessment training at all within a five year period (see Section 3.5.1). This implies that teachers may lack the required level of assessment literacy for successful implementation of assessment in their classroom (see Section 2.4.1). Effective measurement of the learning progress requires teachers to demonstrate a required level of assessment literacy to be able to complete their assessment obligations including to "design, develop, maintain or evaluate, large-scale standardized and/or classroom based tests..." (Fulcher, 2012, p. 125). The apparent teachers' limited knowledge of classroom assessment coupled with their lack of assessment resources to guide them in their assessment practices can explain why most of them preferred to use *paper and pencil* forms.

6. 1.2 Assessment of learning and a focus on linguistic skills

Results on sub-research question 1.2 regarding the purpose and timing of the teachers' use of assessment forms suggested that the teachers' assessed for summative purposes (assessment of learning). Their priority seemed to be gathering evidence of students' learning achievement for

promotional and reporting purposes. Research findings pointing to the teachers' focus on assessment of learning are not uncommon in research studies on classroom assessment (e.g. Mekonnen, 2014; Oz, 2014; Tante, 2010). A study by Tante (2010) examined the purpose of teachers' assessment in English-speaking primary schools in Cameroon and found that teachers' assessment mainly served a summative function with very little attention given to assessment for learning. He found that the majority of the 262 surveyed teachers from grade 1 and 2 of Cameroonian primary schools teachers' assessment was characterized by the common use of summative assessment usually provided for promotional and reporting purposes. Teachers indicated that they assessed mainly to measure the students' learning achievement using promotion exams and end of term exams.

In the same way, teachers in the current study were asked to indicate the purposes for which they assessed their students. Information on why teachers assessed their students was sought to explore the teachers' understanding of assessment and to help examine whether the selection of assessment forms was motivated by the objectives that teachers wanted to achieve in their assessment. Determining the purpose of assessment prior to the implementation can help teachers gather complete and relevant information leading to effective decision taking that benefits both the teacher and the students as Anderson (2003) stresses:

Before deciding how to assess the students, teachers must determine the purpose of the assessment, the assessment information that is needed to accomplish the purpose and the timing of the assessment. The primary purpose of assessment is to gather the information that teachers need to make sound, defensible decisions [...] that can substantially improve a teacher's effectiveness in working with his or her students. (p. 47)

As discussed in section 2.2.2, teachers' assessment may have different aims such as measuring what students acquired from the lessons, commonly known as assessment of learning and summative in nature (Earl, 2010; Leung and Mohan, 2004; WNCP, 2006). Assessment can also be formative in nature, used to help teachers reflect on their teaching and adjust their instruction, to inform students of their learning goals and to provide them with constructive feedback for improvement (Stiggins, 2002; Black et al., 2004). Teachers may also assess to help students become self-motivating through self-assessment, to allow them to reflect on their own learning, by realizing their own strengths and weaknesses (WNCP, 2006, Earl, 2003).

The questionnaire survey results on the teachers' purpose of assessment (Table 4.5) suggest that teachers assessed for both formative and summative purposes. It can also be seen in Table 4.5 that providing feedback to students for improvement, which is common in formative assessment, was highly rated at the top of other purpose statements in the teachers' questionnaire. The results appear to show that the teachers' primary purpose of assessment was to help improve their instruction and guide their students on how to make improvements towards their learning goals. However, the results from the analysis of the qualitative data from classroom observation and interviews seem to suggest otherwise, showing that teachers were inclined to use assessment of learning rather than assessment for learning. This illustrates a contradiction that often exists between teachers' reported beliefs and what they actually do in their classroom (Mekonnen, 2014; Qassim, 2008). One of the possible explanations of these discrepancies may be related the challenges that teachers face when implementing assessment in spite of their willingness and commitment to use formative assessment. In his mixed methods study on the classroom assessment practices of twenty one EFL secondary school teachers in Ethiopia, Mekonnen (2014) found that teachers' assessment beliefs were not reflected in their classroom assessment practices. He recorded a number of factors cited by the respondents themselves on why it was practically impossible to use some assessment forms in their classrooms. Some of those factors included the large class size where assessment activities that were perceived as time consuming were barely used. The respondents also mentioned the lack of motivation on the side of the students and lack of convenient facilities at their schools. The mismatch between teachers' stated beliefs and their observed practices in the current study may be explained by the same factors as in Mekonnen's study. Teachers in the surveyed schools taught a relatively large number of students in their classrooms (up to sixty students in some cases) and reported that access to appropriate assessment resources was one of their biggest challenges.

It could also be that the observed disparity between findings from different data sets was a result of the limited assessment literacy of teachers (see Section 6.2.1) who may have wrongly believed that they were using formative assessment despite the evident summative nature of their assessment practices. In her study on the teachers' beliefs and practices about assessment, Sawafi's (2014) analysis of qualitative and quantitative data from 312 EFL

teachers showed that a mismatch existed between the teachers' beliefs and their actual practices in the classroom. She observed that such a mismatch was as a result of the teachers' limited knowledge and skills in using different assessment methods among other factors. In the current investigation, it was noted during interviews that some teachers may have falsely believed that they provided effective feedback that helped their students to make progress toward their learning goals. Although providing students with feedback was stressed by teachers during interview as their first purpose of assessment, data from the students' focus groups and from classroom observations indicated that the feedback provided by teachers during classroom assessment was usually based on marks and rarely suggested ways for improvement. In addition, the use of marks as the main form of feedback was common among teachers. This is arguably incompatible with formative assessment since teachers made very little effort to go back to the marks and explain to the students what those marks actually meant in terms of progress towards their learning objectives. Instead, some teachers explained that they used the students' marks as an indicator of achievement of their own teaching objectives (e.g. Teacher JS and SL) and no reference was made as to how the students were assisted in understanding the real meaning of their scores.

Teachers often provided whole class feedback consisting mainly of answers of the questions and no further comments were given in the form of suggestions for improvement. In most cases, the grammar based assignments were marked and feedback consisted of distributing the marks to students and correcting the assignment in the classroom. While the marks were considered by students as an indicator of their level of performance in assignments, no further suggestions were provided to them as to what they needed to do to improve. In other words, the feedback that teachers provided did not fit the three-level model of 'feed up', 'feedback' and 'feed forward' (Hattie and Timperley, 2007). According to Hattie and Timperley, effective feedback answers three questions of (1) *Where am I going? (Feed up)* consisting of re-examining the learning goals, (2) *How am I going? (Feedback)* about the progress being made towards the learning goals and (3) *Where do I go next? (Feed forward)* which looks at the activities that need to be done to make better progress. These questions are expected to help teachers provide important information to their students with regard to the entire learning process and not only providing answers of what is right or wrong. The results of

the analysis showed that the teachers' feedback was more aligned with the latter, relying more on traditional assessment that puts more emphasis on assessment of linguistic skills and relies more on assessment of learning than on assessment for learning (see section 2.2.2). It did not equate with constructivist view of feedback resting on the encouragement of both the teacher and learner's engagement in the assessment process (Title, 1994). From the constructivist perspective, the learner's participation in the assessment process fosters the learner's learning ownership and maximises his or her use of assessment information as Title (1994) explains:

A cognitive constructivist perspective [...] suggests that teachers and learners construct schemas or integrate representations from assessments into existing views of the self, of teaching and learning, and of the curriculum, broadly construed. These interpretations include knowledge and beliefs and may also result in intents to use and actual use of assessments. (p.151)

Nonetheless, the students were not given the opportunity to ask questions and teachers often carried on with their lesson in spite of some students failing the assignment.

It is important to note, however, that these results are based on somewhat limited data and the data may not be representative of some of the teachers' usual assessment practices. As explained earlier, each classroom was observed for approximately 50 minutes. Although this was the normal duration of a typical one hour lesson, longer observation may have provided more information on the assessment practices of the observed teachers. In addition, all the seven observed classrooms were from day schools which were often disadvantaged in terms of schools facilities and sometimes had less experienced and less trained teachers as highly qualified teachers preferred to go to boarding schools where they got higher pay (Williams, 2015). It was not possible to conduct classroom observation in boarding schools due to the requirement by the local ethics committee to obtain written parent's consent for every participating student. The fact that the students in boarding schools often came from different parts of the country made it practically impossible to reach every student's parents. Only the students in day schools were able to have their parents' consent forms signed before they could be observed. For this reason therefore, the collected data may not be representative of the reality in all the schools in the lower level of the secondary schools in Rwanda and may not reflect the whole reality about assessment literacy of teachers at this educational level.

6.1.3 Teachers' feedback centred on Marks

The teachers' methods of providing feedback to students were investigated guided by sub-research question 1.4 to gather data on how students were informed about their progress towards their learning objectives. As an important part of assessment, feedback informs students about their level of progress toward the target learning targets and gives them the opportunity to discuss their weaknesses and to decide on how to address their difficulties (Black and Wiliam, 1998b; Hattie and Timperley, 2007). By providing merited positive feedback and persuading students that they can achieve high, teachers develop and strengthen the students' beliefs of self-efficaciousness (Mills and Pajares, 2007; Schunk, 1982). In the current study, data on teachers' methods was gathered to examine whether the provided feedback played a significant role in enhancing the students' self-efficacy. The results indicated that the teachers' feedback was based on verbal comments intended for the whole class and particularly centred on marks. Marks seemed to be the common methods of providing feedback where assessment tasks were either marked for final grade or were set up for motivational purposes.

The role of marks and grades in classroom assessment is often discussed in literature and studies have shown that their use as part of teachers' feedback is still common in spite of recent calls for use of more informative feedback (Black and William, 2009; Lee, 2009). In a study conducted by Lee (2009), the analysis of texts from 206 EFL secondary school teachers in Hong Kong to identify their feedback practices found that teachers still used marks and grades even when they 'did not have faith' in such feedback. His interview with some of these teachers revealed that although they knew that feedback in the form of scores could divert the students' attention, scoring the students' work was necessary in order to obtain the required final grade. This illustrates how contextual factors related to policies and socio-political issues can affect the teachers' practices. Teachers in test-oriented pedagogies and in high stakes tests contexts often use marks and grades as evidence of learners' achievement and often pay little attention on other forms of feedback (Shohamy, 2005). However, most literature on classroom assessment has warned against the possible adverse effects of using marks as a method of providing feedback (Black et al. 2002, Black and Wiliam, 2009; Rust, 2002). As Black et al.

(2002) argue, reliance on marks or grades may erode the students' interest in feedback provided in the form of comments:

When giving students feedback on both oral and written work, it is the nature, rather than the amount, of commentary that is critical. [...] while student learning can be advanced by feedback through comments, the giving of numerical scores or grades has a negative effect in that students ignore comments when marks are also given. (p. 13).

Both oral and written feedback was provided in the form of marks, corrections of questions or short comments of praise to name a few. The results also indicate that marks and oral comments provided to the whole class were the most frequently used forms of feedback. On the one hand, oral feedback in the form of short comments was often used to give praise to individual students for correct answers to teacher's display questions. On the other hand, marks were usually accompanied by correction of assignments with answers being written on the chalkboard for the students to record them in their notebooks. These categories of feedback were used for both in-classroom and take-home assessment tasks.

Given the relatively high classroom-student ratio average of 35:1 in Rwanda Lower Secondary Schools (MINEDUC, 2015b), and the pressure to cover all the curriculum content, the teachers' frequent use of verbal and whole class feedback was anticipated. Four of the seven interviewed teachers (Teachers MR, VN, JZ & SL) argued that they preferred to use feedback addressed to the whole class because individualised feedback was considered to be time consuming for them. However, although it was anticipated that some teachers would report not being able to provide regular individualised feedback to all students due to factors such as large class size, some of the teachers' stated reasons for not providing individualised feedback were unexpected. For example Teacher MR pointed out that she never used written individual feedback on fear that addressing individual students could lead to adverse consequences which might result from the misinterpretation of her comments. As she explained, this was only motivated by her past experience of seeing her former colleague accused of harassment because of his comments that students often tagged as negative. However, the fact that Teacher VN reported that he successfully used individual written feedback that Teacher MR considered as inappropriate while both of them taught at the same school could be viewed as additional evidence that teachers did not have opportunities to share their assessment experiences (See

Section 6.2.1). It also seemed to show that teachers likely lacked support of their school administration to help them work together to adopt appropriate methods that would be perceived as helpful by learners instead of being seen as potential source of harassment.

Some teachers also indicated that they were overly sensitive to their students' feelings and avoided feedback that could confuse or be misinterpreted. For example, Teacher MR and Teacher SL both seemed to hold the belief that feedback that was not meant to praise would be considered by students as undesirable. As result, they chose to use marks from objectively marked tests. As discussed in the previous section, these beliefs seemed to be in conflict with the basics of formative assessment. Although avoiding 'negative' feedback could affect the students' self-beliefs by increasing their perceived self-esteem and confidence (Oroujlou and Vahedi, 2011; Pui, 2010), the fact that Teachers MR and SL avoided any source of conflict with the students by focusing on positive feedback only i.e feedback relating to success, could have other unwanted effects. Students may develop high self concept and always aim to protect their positive self-evaluation instead of engaging in their tasks for real learning to take place (Kluger and DeNisi, as cited in Baadte and Schnotz, 2013). It appeared that these students failed to get the benefit of the constructive feedback as they sought to avoid negative feedback which they considered as a threat to their self-concept.

Teachers MR and SL's claims could be regarded as further evidence of some teachers' limited assessment literacy as they seemed unable to formulate suitable and constructive feedback that could not necessarily be viewed by learners as inappropriate. However, it can also be argued that by favouring positive feedback, learners also demonstrated limited knowledge about effective feedback and limited assessment literacy in general. The students' assessment literacy is defined as "students' understanding of the rules surrounding assessment in their course context, their use of assessment tasks to monitor or further their learning, and their ability to work with the guidelines on standards in their context to produce work of a predictable standard" (Smith et al., 2013, p.46). Smith et al. argue that students need to know the purpose of assessment and how it links with their learning goals, develop awareness of assessment process and get the opportunities to practice following assessment guidelines and judging their works against a predefined standard. The findings of the current study suggest that the students were only familiar with marks as their usual feedback from teachers. They also

demonstrated a limited awareness of the potentials that a range of assessment activities could provide them and they were rarely consulted by teachers for the design of assessment tasks. In addition, the fact that the surveyed students were at a young age, including twelve year olds and below, may suggest that some of them associated positive feedback with high ability, thus demonstrating a less advanced ability judgment which is common in young learners (Barker and Graham, 1987). This may explain why there was no data suggesting that students wanted to see a more honest appraisal of their work.

The process of providing feedback as described by teachers and students reflected a more or less established tradition based on marking apparently stemming from the education system that seems to encourage competition and outcome based rather than process based assessment (See Section 6.2.2). Teachers also demonstrated an overall preference for positive and collective feedback to individual feedback. They also avoided feedback that might be deemed ‘negative’. Like teacher VN, some teachers acknowledged the ineffectiveness of these methods of providing feedback but continued to use them due to other influencing factors as described earlier (e.g. large class size). Although some teachers reported that they used some other forms of feedback (e.g. summoning a student for a face-to-face advisory talk following the student’s poor performance on a marked assessment), it seemed that such feedback were less frequent. Teachers appeared to stick to the method (use of marks) that students were familiar with. Generally speaking, the students were satisfied with the teachers’ feedback and grading (see Table 4.10). This was not surprising given that the majority of the respondents had achieved high in relation to English subject. While a score less than 50 out of 100 was generally considered as a fail, more than a half of the surveyed students had a score higher than 60 out of 100 in English subject for the previous term (mean= 64.84, mode:60, median:65). This might justify their feeling of satisfaction with their teachers’ feedback. Nonetheless, the fact that the students were only familiar with marks as a common feedback and less aware of other possible feedback options and the potentialities that they could offer might have played a crucial role in determining the students’ perceptions of their teachers’ feedback.

Overall, based on the findings of the current study, it might be fair to say that marks-based feedback used by teachers in their classroom fell short of information that the students needed in order to know about their learning progress and what needed to be done to bring

about improvement. For feedback to be effective, it has to take into account the teaching and learning objectives and reflect on the progress made to suggest ways for improvement (Alvarez et al., 2014; Hattie and Timperley, 2007). As explained earlier in Section 6.1.2, it is vital to recall the three levels of effective feedback suggested by Hattie and Timperley (2007):

Effective feedback must answer three major questions asked by a teacher and/or by a student: Where am I going? (What are the goals?), How am I going? (What progress is being made toward the goal?), and Where to next? (What activities need to be undertaken to make better progress?) These questions correspond to notions of feed up, feed back, and feed forward. (p.86)

According to Hattie and Timperley, teachers need to base their feedback on these three notions of ‘feed up’, ‘feed back’ and ‘feed forward’ in order to provide information that helps learners to understand their learning goals, identify their progress toward those goals and suggest ways to attain more.

Drawing on the results of the data analysis, it seemed that the students were satisfaction with their teachers’ feedback. This can be explained by the fact that the students often expected and indeed usually got satisfactory marks from assessment tasks as the main feedback. Despite the students’ beliefs that their teachers’ feedback was informative, it can be argued that such feedback did not aim to help the students develop awareness of their learning progress. However, it is also worthwhile recalling that data on the teachers’ methods of providing feedback was based on teachers’ and students’ reporting. This is significant as the method that was used might not be entirely appropriate for collecting reliable information. In fact, the use of document analysis such as the teachers’ assessment records, or the students’ assignment sheets displaying the teachers’ comments might have provided more reliable data. Nonetheless, the results from the analysis of data collected in the current study highlights the importance of feedback especially in shaping the students perceptions about assessment as a whole.

6.2 Factors influencing the teachers classroom assessment practices

The discussion of findings above provides some insights into the possible factors that seemed to influence the teachers’ choice of assessment forms. Based on the current results, it shows

that the teachers' preferences in classroom assessment forms may be related to factors such as the teachers' own assessment beliefs, their levels of assessment literacy and the insufficiency of assessment resources. Finding appropriate and sufficient materials was cited by teachers as one of their main challenges while the results on their in-service training in classroom assessment (see Figure 3.7) suggest that they lacked assessment knowledge and skills necessary to deal with such challenges. There also seemed to be no support available at many schools to help teachers acquire assessment knowledge and skills required for the effective use of different assessment methods. In addition, the findings seem to suggest that the national top-down system of assessment in Rwanda could be another influencing factor where washback effects of the national examinations on the classroom assessment practices were alluded to by the interviewed teachers. The following section aims discusses these factors further.

6.2.1 Scarcity of assessment resources and teachers' limited assessment literacy

The results from the analysis of data on sub-research question 1.3 regarding the sources of assessment activities that teachers used in their classrooms suggest that due to the shortage of assessment materials aligned with the curricula and adapted to their learners' levels, many teachers were compelled to design their own assessment materials or look for external alternatives. For example, the results from both the qualitative and quantitative data revealed that many teachers reported using their own notebooks from college for assessment activities while others used tasks from the official course textbooks. However, the majority of the interviewed teachers (66%) stated that they designed and used their own assessment tasks (see Section 4.1.3.4). In some cases, teachers designed assessment activities as they taught (teacher FR and SL). These often consisted of short answer questions that were usually grammar based that teachers used to measure the students learning at the end of lessons. Some teachers reported that the insufficiency of appropriate materials sometimes compelled them to look for other sources including internet despite its limited availability at many schools.

Access to adequate resources and well prepared assessment materials can be essential for assessment effectiveness in contexts where teachers are not well trained in designing their own assessment materials. Teachers well trained in assessment with high teaching experience

often develop their own items (Chen et al. 2004). This is done particularly when they are not able to find readymade assessment materials or when they perceive that the available materials are not suitable for their classroom assessment (Shim, 2009). The current study reports similar situations as explained by Teacher SL who pointed out that her five year experience had provided her with the ability to design assessment tasks even during the course of the lesson. Although up to 70% of the surveyed teachers indicated that they had at least 5 years of teaching experience, it can still be argued that the majority of them possessed limited knowledge and skills on how to develop appropriate assessment items given that only a small percentage of them had been trained in classroom-based assessment design. Studies have shown that the teachers' level of assessment literacy is one of the key factors that influence their choice of assessment practices (Yang, 2008; Zhang and Burry-Stock, 2003).

Based on the findings on the sources of teachers' assessment materials, it seems that the paucity of assessment resources at schools urged many teachers to design their own assessment activities. The majority of the interviewed teachers indicated that it was hard for them to get access to internet or published materials both at schools and outside the schools. Some teachers also opted to use question items from their own notes taken during their studies at college or model their assessment activities to those used by their former teachers in the college and hence promoting the grammar focused approach to assessment. The absence or insufficiency of appropriate audio-visual equipment such as audio players could explain why assessment of listening skills was rated at a lowest level (See section 1.3 for the classroom context). With the lack of ready listening tasks and being unable to design them on their own, some teachers might feel that they were unable to effectively assess this receptive skill. In addition, the fact that the National Ordinary Level Examination did not include listening tasks might discourage teachers to train and assess their students on listening (see Section 2.1.3.2 and Section 6.2.2 for washback effects).

It is also important to note that teachers in this study demonstrated that they rarely developed items together with their colleagues, which can arguably raise further questions on the validity of assessment items developed by individual teachers on their own (see Section 2.4.1 on validity). For example, the findings of this study showed that Teacher MR and Teacher

VN at school C (one of the 7 schools where observation was conducted) had opposing views on how to provide feedback appropriately and the role that some feedback could have on learners. This proved that there may be no collaboration between these two teachers which led to each of them sticking to their own preferences. With no opportunity to meet and share their assessment experiences, teachers at school C were not able to learn from each other regarding assessment alternatives. Teachers' collaboration during assessment helps them increase the validity and reliability of their assessment by sharing experience and learning from each other (Akyel, 2000; Allal and Lopez, 2014). It allows for "consistency and comparability of outcomes for reporting and accountability purposes, improved quality of assessment (in support of learning), increased coherence of assessment practices and transparency of assessment practices" (Allal and Lopez, 2014, p. 125). Put in a more general perspective, Mann (2005) argues that the teachers' "focus groups" or "collaborative groups" can be vital for their professional development as they offer opportunities for collaborative peer-discussions:

What is central about such groups is that understandings are constructed through talk. It is not a matter of simply sharing and transferring information, rather than arguments, understandings, clarifications, and interpretations are constructed through spontaneous conversation with other professionals. (p. 111)

Teachers who participated in the current study had a wide range of professional experiences, from less than a year of teaching experience to up to twenty years of experience in English language teaching. In many cases, the experienced or highly qualified teachers taught at the same school as the inexperienced or the less qualified colleague. In these cases, collaboration by group discussion could have been a good opportunity to share materials and widen a shared understanding of assessment particularly for the benefit of novice teachers. In addition to discussions, teachers could also learn from each other through 'co-teaching'. As another form of collaboration that has been used in some EFL contexts, 'Co-teaching' consists of two teachers (usually a native speaker and a non native speaker) working together to deliver a lesson (Liu, 2008; Stepp-Greany, 2004). Although this could be very useful for the teachers, it did not seem easily feasible as it would require important changes of the teaching timetables. Teachers often had heavy teaching workloads and colleagues at the same school usually taught their respective classes at the same time slots. It appears however, that the 'teachers'

collaborative groups' as Mann (2005) suggests, could be more practical and convenient seeing that they could be carried out at the teachers' preferred venues and at their convenient time.

6.2.2 The Washback effects

The analysis of different data sets in this study also reveals that there may be some washback effects to the teachers' classroom assessment practices given that the national examination taken by students at the end of lower secondary school is considered as a high stakes test by both teachers and students. Students want to score high on this test in order to get admission to upper level of secondary school and as a result, teachers put much focus on preparing their students on how to take this test in order to get high grades. A high stakes test can have the power to determine what happens in the classroom with either bad effects (negative washback) if students and teachers are negatively affected or with good effects (positive washback) if the test makes "teachers and learners do 'good' things they would not otherwise do" (Alderson and Wall, 1993, p. 117). The results of this study suggest that the administration of the national examination at the end of year three of the lower level of secondary schools in Rwanda affected the teachers' preference of classroom assessment forms and the content focus of their assessment. Teachers of F3 reported that they used past papers of the national examinations for their classroom assessment to familiarise their students with the national examination thus demonstrating the immediate effect of this test to the classroom assessment level.

The surveyed teachers confirmed using past national examination papers and focusing on the same type of question items. It appeared that teachers chose the assessment forms to use depending on the learning content being assessed at the national level and they focused on grammar and text comprehension which were the two main components of the national English language examination. While this can be seen as a good effect given that the teachers seemed to take the tested areas seriously (Alderson and Wall, 1993), it also had a negative effect given the fact that assessment was only focused on certain parts of the curriculum to the expense of other curricular components. The language curriculum for the lower secondary school contained learning content pertaining to speaking, writing, listening and reading skills in addition to grammar and vocabulary. Despite these curricular components, more attention was given to content that was likely to be assessed in the national examination mainly grammar and reading

comprehension. The results from both quantitative and qualitative data analysis indicated that teachers focused more on grammar than any other content during assessment (see section 4.2.3.3). The focus on grammar in assessment may explain the frequent use of paper and pencil based forms of assessment which are commonly used for vocabulary or grammar related tasks as Rhodes, Rosenbusch and Thompson (1996) highlight:

Most traditional assessments of language skills tend to be discrete point tests, with emphasis on linguistic accuracy, such as grammatical structure and vocabulary. Discrete-point tests tend to focus on single skill areas and evaluate the knowledge of details of the language. Items are typically presented in single sentences or phrases that are unrelated and lack context. (p. 382).

As the National examination did not assess the students' speaking and listening skills, the teachers were less interested in assessing these skills but focused more attention on the grammar centred assessment forms. For example teacher MR reported that she used *short answer* questions or *gap fill* when she wanted to assess the students' knowledge in grammar and preferred to use class presentations for assessing speaking. During classroom observation, it was noted that most classroom talk came from teachers usually explaining a grammar point or asking students to answer grammar related questions (see Table 4.3). In addition, teachers in six out of seven observed classrooms taught and gave assessment on grammar either through teacher oral questioning or through written work done in the classroom or as take-home tasks.

Comments from the interviewed teachers suggest that their use of past exam papers was not necessarily motivated by the perceived relevance of those exam papers to the learning content and learning objectives. The teachers in F3 explained that they used past Ordinary Level National Examination papers mainly to prepare their students for taking the same examination. One implication of this may be that past exam papers were used for 'drilling' purposes where students are primarily trained on how to get the highly demanded grade. This is in contrast with the view that access to past exam papers can be of great importance when they are put to appropriate use or when teachers use them as an inspirational source for development of their own assessment materials that match the content and objectives of the language class (Rhodes, Rosenbusch and Thompson, 1996; Watanabe, 1997). As in the case of teachers' high reliance on paper and pencil forms of assessment, the frequent use of these examination papers also suggest that the teachers focused most on the linguistic skills given that the examination

itself was grammar-based. As discussed earlier, this is part of tests' washback effects. The interviewed teachers referred to the grammar-based National Examination to justify their teaching and assessment practices. They explained that focusing on grammar-based tasks could help them prepare their students for the national examination. It is important to recall that this is a high stakes examination that is taken by F3 students to test their eligibility to continue their studies in upper secondary school (MINEDUC, 2010).

Literature on test washback has indicated that in contexts where high stakes tests are used, teachers often choose their teaching and assessment priorities depending on what is likely to be assessed in the test (e.g Harlen and Crick, 2002; Hsu, 2009; Watanabe, 2004). Hence, it was not unexpected that reading and writing were rated by both the students and teachers as being the most frequently assessed skills considering the important place that they had in the national examination tests. This appeared to be contradicted to some extent by the results in Table 4.7 which show that teachers rated reading as being rarely assessed, this might have been caused by the teachers' misinterpretation of what constitutes a reading task. When the name of the specific reading tasks was used, reading was one of the two top methods of assessment both from students' and teachers' responses (see Tables 4.1 and Table 4.2). Some teachers in the current study expressed their disapproval of some of their assessment practices (Teacher MR and SL) that they considered less effective. However, the top-down assessment system where teachers were expected to help their students get satisfactory scores in the high stakes national examination compelled them to continue with the practices that are common and widely accepted in their context. This matches what Davison (2004) describes as the 'teacher's beliefs' about others' judgment against their behaviour:

If teachers believe that those important to them in their professional community will approve a certain practice, then they will be predisposed to carry out that practice, even if it is not acceptable in the wider community, and vice versa. Other contextual factors include teachers' beliefs about the social, institutional and cultural context of their assessment practice, including the purpose of assessment, its relation to learning and teaching, the role of the teacher in relation to assessment, and the teachers' preexisting beliefs about the students and texts they are assessing. (p. 308)

The interviewed teachers explained that the pressure to have their students perform well on the national examination were some of the key factors behind the choice of assessment forms that

they used. As explained above, this was the case particularly for teachers of F3 (e.g. Teacher VN and SL) who often used mock exams from past National Ordinary Level Examination. Teachers expressed their feeling of pressure not only for completing the entire curriculum content but also ensuring that their students had the opportunity to practice the techniques and language used on the national examinations.

Results from this study provide supporting evidence to the findings of other studies which have revealed that high stakes examinations at both local and national levels can influence the teaching and learning process at the classroom level (Alderson and Wall 1993; Hsu, 2009; Shohamy, 1997; Wall 1997). It was evident from the teachers' comments that the national examination washback existed and affected their decisions on what to use for classroom assessment. The fact that teachers spent a significant part of their assessment preparing their students to sit for the national exams highlights the existence of the test's washback effect on classroom practices, viewed as negative since it diverted teachers from focusing on the curriculum content (Bailey, 1999; Nkosana, 2008; Azadi and Gholami, 2013). Instead of focusing on the teaching and learning content in the curriculum which would allow opportunity to use different assessment forms, teachers preferred to narrow the curriculum content down and focus on what would come on the test, hence using assessment forms appropriate to their grammar based lessons.

While teachers in F2 had grammar as their central teaching and assessment focus, teachers in F3 also taught and assessed their students on productive skills such as reading and summary writing. Emphasizing and in some cases expanding the content assessed on tests and contained in the teaching and learning objectives is viewed by some as positive test washback (Brown and Hudson, 1998). However, in the context of the current study, the negative test washback seems to be more significant than positive washback. As the test itself is largely based on grammar, it encouraged teachers to rely more on assessment forms that are appropriate to grammar to the expense of performance based assessment.

6.3 The students' perceptions about the teachers' assessment practices

This section discusses findings on sub-research question 1.5. The analysis of the students' perceptions of their teachers' assessment practices was carried out to examine their influence on the students' self-efficacy. Previous research studies have investigated the students' perceptions of classroom assessment and have reported the influence of the students' perceptions on their academic self-efficacy (e.g. Alkharusi et al. 2014, Dorman et al., 2006). In their study of the relationship between the students' perceptions of classroom assessment and academic self-efficacy of 1457 Omani secondary school students, Alkharusi et al. (2014) found that academic self-efficacy was significantly and positively influenced by the students' perceptions of assessment tasks. Using multilevel regression techniques to determine the strength of association between these two variables, Alkharusi et al. found that students' perceptions of congruence with planned learning, authenticity, transparency, and diversity of assessment tasks all had influence on the students' academic self-efficacy.

Similar findings were also obtained in Dorman et al.'s (2006) study that investigated 449 Australian secondary school students on their perceptions of classroom assessment and their relationship with academic efficacy. The results from multiple regression analysis indicated that the students' perceptions were significant predictors of the students' self-efficacy. In the case of the current study, understanding how students perceived their teachers' assessment practices was expected to provide some insights into the basis upon which the students judged their self-efficacy for communicating in English (i.e speaking, writing, reading and listening). Generally speaking, the results of the study indicate that the surveyed students expressed positive perceptions toward their teachers' assessment practices in terms of assessment *congruence* with the learning content, the teachers' *fairness in grading*, the *frequency* of assessment that they did in the classroom, the way in which teachers *diversified their assessment* methods and the frequency of the teachers' *feedback* as a whole (see section 4.3).

During interviews, students in all four focus groups stated that they often did written assessment in the classroom and were given regular homework most of which was marked.

This may explain their high rating on assessment *frequency* and *diversity* as can be seen in Table 4.10. It also appeared that the students' reported satisfaction with *frequency* and *diversity* of assessment tasks was influenced by the students' limited assessment literacy. Seemingly unaware of the different forms of assessment that can be used for classroom assessment, the students might have falsely believed that they had enough and effective classroom assessment experience. It is highly likely that the students' perceptions might have been different if their assessment literacy level was higher. This was the case in a study of the students' perceptions of classroom assessment by Mussawy (2009). In this study that investigated the perceptions of 28 Afghan students from a higher learning institution in Afghanistan, Mussawy (2009)'s description of comments made by students highlights that they were aware of different forms that their teacher could use for their classroom assessment:

Students' responses show that they realized the forms of assessment that are helpful in their learning. [...] Generally, the current practices of classroom assessment seemed dissatisfying for some students given the dominance of traditional forms, assessing students at the end of the semester. Using assessment at the end of the semester leaves an instructor very little chance to use the assessment results to improve instruction and students' learning. Overall, there seems an understanding of various forms of assessment by students (p.64)

As teachers failed to use the assessment forms that students expected to see in their classroom, the students expressed their dissatisfaction with assessment as a whole. Their apparent assessment literacy enabled them to realise that appropriate assessment forms were not used.

In the current study, the students also reported highly positive perceptions of their teachers' fairness in marking. The prevalent use of traditional forms of assessment (e.g sentence completion, true/false questions) with most of assessment assignments invariably focused on grammar, it was easy for the students to understand how their teachers marked their work as the marks were objectively assigned for either wrong or right answer. The lack of assessment that allows for subjective answers in the form of open ended questions or performance assessment tasks where answers are not necessarily wrong or right could have helped to avoid the students' confusion and complaints about marks or the marking process. This might have influenced their satisfaction with the teachers' *fairness in grading*. Interviewed students explained that the usual feedback that they were familiar with was marks

and the correction of assignment questions on the chalkboard. The fact that the students expected marks and the collective correction of the questions after completing assessments as the only *feedback* from teachers could explain why they reported being satisfied with how they received feedback. It seemed that students were not aware of other methods of providing feedback that their teachers could use, which also relates to their level of assessment literacy discussed earlier.

Although students were positive about the *authenticity, frequency* and *diversity* of assessment as well as teachers' *grading fairness*, overall *feedback* and *congruence* with learning content, data from both the questionnaire survey and student focus group interviews indicated that students were not satisfied with *transparency* of assessment aims and were rarely involved in the decision making process about assessment (*student consultation*) (see Table 4.10). *Transparency* of assessment refers to "The extent to which the purposes and forms of assessment tasks are well defined and clear to the learner" (Dorman et al., 2006). Transparency can be increased when teachers explain the criteria used to score the students' work (Cheng and Wang, 2007). During interviews, the students stated that they often did not understand what the assessment aims were. This may seem contradictory given the fact that students had reported higher grades in classroom assessment. However, the students explained that they achieved high on assignments because the format of most assessment activities was unchanged and that they usually knew how to answer (see section 4.4, S2 from Group B).

Students could also have good grades without necessarily understanding what they had to do on assessment during group work. It was observed during classroom observation that where group work was given to the students during assessment, only a limited number of students were actively engaged in the work, completing the task while others looked on passively. Although some teachers (e.g Teacher SL and Teacher JS) monitored from group to group, their main concern was to verify if the group as a whole had arrived at the final answer. This was itself product oriented, usually in the form of short right or wrong answers. There was no mechanism set before the tasks to ensure that every student participated. Once the group work was graded, each student in each group was assigned the same group grade while it was rather a work by a very limited number of group members. The students were satisfied with how the teachers marked but they indicated that assessment tasks lacked *transparency* that

should be the centre of all assessment practices if formative assessment is to take place (Torrance and Pryor, 2001). Greater *transparency* in assessment tasks also helps the students to understand what constitutes a high-quality work (Topping, 1998).

In the context of the current study, the lack of *transparency* as reported by the respondents might be due to the non use of rubrics among other causes (see Section 2.2.5). Teachers demonstrated that they had limited knowledge of the importance of using rubrics in assessment and only one of the seven interviewed teachers had used them. When appropriate rubrics are well used, the learning targets become clear and the students become aware of what their teachers expect them to do (Andrade et al., 2009). Rubrics are also important for the development of the students' self-assessment abilities as Jonsson (2014) explains:

It is indeed possible to convey expectations to students through the use of rubrics, in the sense that students not only appreciate the efforts to make assessment criteria transparent, but may also use the criteria in order to support and self-assess their performance.

Given that teachers used paper and pencil based assessment and focused more on assessing linguistic skills to the detriment of performance skills, the use of rubrics in their classroom may arguably seem extraneous. As discussed earlier in Section 2.2.5, there has been a growing assumption that rubrics are more relevant for performance based assessment than paper and pencil based assessment (Andrade, 2000; Mabry, 1999; Outeiral, 2014). However, the communication of the learning goals and ways to achieve those learning goals as well as the criteria used to assess achievement is as important in paper and pencil based assessment as it is in performance based assessment contexts. Nonetheless, this did not seem to be the case in the surveyed classrooms. The students' comments indicated that they relied on their familiarity with the teachers' assessment methods suggesting that the students did not receive enough clear and meaningful explanations from the teachers regarding assessment expectations.

Another way that transparency could have been increased is by use of exemplars. These are "examples, typical specimens or model answers for an assessment. They demonstrate to students the type and level of performance expected for an assessment task" (Shapland, 2011, Model exemplars). The use of exemplars can also help the students to develop their self-assessment skills as they analyse real marked or unmarked assignments from previous

assessment (Stefani, 1998). The results from analysis of data suggest that teachers did not use exemplars in their classroom assessment for reasons that may include teachers' apparent low familiarity with their use, lack of resources and time constraints. In fact, one of the drawbacks of using exemplars is the large amount of time needed for construction and sourcing of appropriate specimen of works that students can analyse to understand criteria and standards of a quality work (Newlyn, 2013).

6.4 High self-efficacy for reading and speaking

Sub-research question 2.1 aimed to explore the students' self-efficacy levels in each of the four English language skills. The students were asked to indicate the extent to which they could perform tasks in the four English language skills of speaking, writing, reading and listening. The results showed that reading and speaking were the top two skills in which students felt more self-efficacious (see Table 5.2). These results match findings from a study by Genç, Kuluşaklı and Aydın (2016). They used a questionnaire survey to investigate the perceived self-efficacy of 210 EFL Turkish undergraduate students in the four English language skills and found that students perceived themselves as being more self-efficacious in reading and speaking. Another similar survey by Wang et al. (2013) investigated the self-efficacy of 167 Korean EFL college students and found that they reported higher self-efficacy in speaking and reading than in listening and writing skills. The studies by Genç et al. (2016) and Wang et al. (2013) both showed that students reported lowest self-efficacy in listening. The same results were obtained in the current study where Listening was the skill in which most students did not feel capable of successfully completing tasks.

Wang et al (2013) explain that students' perceived low self-efficacy in listening may be influenced by the extreme mental exertion required for the listening process which also causes feeling of anxiety among students. While the influence of perceived anxiety for listening may be a factor in the current study, reports of low students' self-efficacy in listening was not surprising considering that both students and teachers had reported rare use of listening tasks in the classroom. The students did not therefore have evidence such as marks, grades or the teachers' feedback on which they could base their confidence. Nonetheless, having speaking as the top skill in which most students felt self-efficacious was unexpected because speaking tasks

were not among the most frequently used in the classroom compared to grammar related tasks or tasks focused on reading. However, it was noted that both teachers and students considered successful everyday teacher-student interaction in English in the classroom as comparable to assessment of speaking and therefore indicative of the students' capability to speak English well. They also argued that speaking was part of their writing (e.g. during dictation) and reading (e.g. students reading aloud). It may be true to a certain extent, therefore, that the high frequency of teacher-student oral interaction as noted during classroom observation (see Table 4.3) and reiterated during interviews, may have served as evidence to some students that they could successfully accomplish speaking tasks.

It may be argued that the students' perceived self-efficacy was in fact not based on realistic facts given that the classroom teacher-student interaction was always teacher-led and dominated by teacher display questions (see Table 4.3). Although the respondents' comments seemed to imply that they could perform the tasks because they did them in the classroom, data from the classroom observation and from most interviewed teachers suggested that real performance tasks that would provide real demonstration of the students' abilities in speaking or reading were only seldom used. This appears to reemphasise the issue of the students not being given opportunities to practice real life like communication skills and therefore not having full understanding of their actual abilities to perform the stated tasks in English. Nonetheless, the high self-efficacy levels in reading and speaking may have been influenced by some other factors such as grades, age or family background. These seem to greatly outweigh the influence of assessment tasks as was also evidenced by the results from quantitative data analysis (see Section 5.1). In fact, the statistical analysis to test the significance of self-efficacy differences between different groups of students based on demographic and school context factors indicated self-efficacy varied among students as a function of several factors.

It is evident that students may have overstated their self-efficacy for using English proficiently in speaking, listening, reading and writing. This seems inconsistent with the perceptible consensus in the Rwandan academic literature that secondary school students and university graduates alike lack a decent level of English proficiency. It has been often argued that the perceived lack of communicative competence inhibits the students' enthusiasm to

speak English for fear of being negatively judged (See section 1.4.3). Therefore, the current findings of high self-efficacy not only calls for a further analytical look into factors influencing the students' overconfidence, but also highlights the need for appropriate practical measures to help the students understand the real level of their abilities as Pajares (1996) clarifies:

Efforts to lower students' efficacy percepts or interventions designed to raise already overconfident beliefs should be discouraged, but improving students' calibration—the accuracy of their self-perceptions—will require helping them to better understand what they know and do not know so that they may more effectively deploy appropriate cognitive strategies as they perform a task. The challenge is to accomplish this without lowering confidence and optimism. (p.565)

The 'students' calibration' advocated by Pajares is defined by Phakiti (2006) as the “perfect relationship between confidence in performance and actual performance” (p. 36). He also draws attention to the importance of feedback that helps students to match their confidence with actual performance without undermining their self-efficacy. Phakiti (2006) argues that the teachers' feedback can help learners to develop strong 'Internal feedback'. This enables them to have “judgments of success in the task in regards to the desired goals, judgments of the relative productivity of various cognitive processes such as strategies along with expected rates of progress, and positive or negative feelings associated with productivity” (Phakiti, 2006, p. 40). In addition, the use of rubrics is suggested as one of the possible ways to help the students measure their own performance and develop full understanding and awareness of their abilities to perform designated tasks (Panadero and Romero, 2014).

6.5 Factors influencing the students' self-efficacy

Following the discussion above on the reported levels of the students' self-efficacy, this section discusses further the observed differences in the students' self-efficacy as a function of different factors. The focus remains on the findings from the analysis of data that was collected to answer sub research question one on the students' self-efficacy for English speaking, writing, reading and listening. The following sections review the reported self-efficacy and discuss different factors which, as the findings suggest, may have some influence on the students' self-efficacy levels. These include factors related to the teachers' assessment

practices, the school context as well as those related to the demographic attributes of the students.

6.5.1 Students' perceptions and their self-efficacy

As stated earlier, research studies have indicated that the students' self-efficacy is highly related to and can be predicted by the students' perceptions of assessment (Alkharusi et al. 2014; Dorman et al., 2006; Dinther et al. 2014). The current study sought to examine the extent to which the students' self-reported perceptions of classroom assessment were related to their self-efficacy. The findings indicate that the students generally held positive perceptions toward the teachers' classroom assessment practices and six of eight investigated perception variables were all related to the students' self-efficacy. However, only positive perceptions towards *transparency* of assessment aims had a high correlational coefficient in the range between 0.35 and 0.65 (see Table 5.11). Correlational coefficients in this range are statistically significant and may have some predictive ability (Cohen et al. 2007). This highlighted the effect that the level of transparency of assessment aims could have on the students' beliefs about their abilities to use English language. In other words, the students who understood well the aims of assessment tasks and what they were required to do to achieve those aims tended to have higher levels of self-efficacy than students who could not understand what their teacher's assessment aims were. The results of the analysis show that other five out of eight perception variables were all related to the students' self-efficacy in all four English language skills to varying degrees.

In addition, students who reported positive perceptions about teachers' *feedback*, *authenticity* and *congruence* of assessment with learning content and those satisfied with the *frequency* of assessment also reported high self-efficacy levels. These findings mirror those of previous studies that have examined the effect of the students' perceptions on their self-efficacy (e.g. Alkharusi et al. 2014; Dinther et al. 2014; Dorman et al., 2006). In their study on the interaction between the perceptions of competence-based assessment and self-efficacy of 138 Dutch undergraduates enrolled in teacher education programme, Dinther et al. (2014)'s multiple regression analysis indicated that the students' perceptions of feedback and authenticity of assessment did predict the students' self-efficacy. Lending support to these

findings, the results in the current study show that the students' perceptions particularly *congruence* with learning content, authenticity of assessment tasks and transparency of assessment aims, were all significant factors influencing the students' self-efficacy for performing tasks in all the four language skills.

Unlike the findings in Dinther et al. (2014)'s study indicating feedback as one of the most important perceptions affecting self-efficacy, the results displayed in Table 5.11 show that the students' perceptions of feedback in the current study was not particularly highly correlated with their self-efficacy. This may be a result of the nature and limited amount of feedback that the students obtained from their teachers. Interviewed students explained that marks were indicator of their learning success and a common source of information about their learning progress. In addition, the fact that participants in Dinther et al.'s study were students in higher education might have influenced the results. Compared to students in secondary education, students at a tertiary level are likely to have more advanced understanding of feedback and its potentials. From the social cognitive theory perspective (Bandura, 1997), feedback can enhance the students' self-efficacy when they get confirmation of success (mastery experience) and encouragement (social persuasion) from both their classmates and teachers.

On the other hand, no correlation was found between the students' self-efficacy and their perceptions regarding *student consultation* in the assessment decision taking or the teachers' use of *diverse assessment* tasks. These results match the findings from previous studies where no relationship has been found between *student consultation* and self-efficacy (Alkharusi et al. 2014; Dorman et al., 2006). However, contrary to the findings of these studies, no predictive relationship was found between the students' self-efficacy and their perceptions on *assessment tasks diversity*. In this study, the students reported having positive perceptions toward *diversity* of the teachers' assessment tasks. However, the interviewed students cited written classroom tasks and take home assignments which were often grammar focused, as the two main methods of assessment that teachers used for assessment. Nonetheless, the students seemed to view these as sufficient in terms of alternatives hence feeling positive about *diversity*. However, these were very limited range of assessment methods and were not reflective of the complexity of real life communication which might explain why *diversity* of

assessment tasks was not significantly related to the students' self-efficacy beliefs. In other words, the used methods of assessment did not comprise the necessary assessment forms capable of familiarizing the students with various communication tasks that could boost their confidence.

6.5.2 Influence of demographic factors

The analysis of data on the students' self-efficacy also examined other factors that were unrelated to classroom assessment but which were believed to have some effect on the students' self-efficacy. Identification of these factors was expected to help understand the degree of influence that the classroom assessment related factors had on self-efficacy compared to other factors. In this section, the focus is put on the demographic factors namely the students' age and gender. The results presented in section 5.1 indicate that age and gender of students had influence on their self-efficacy. As illustrated in Table 5.3, the comparison of mean ranks from Mann-Whitney test show that male respondents reported higher self-efficacy in all four skills than female respondents. Similarly, the results from Kruskal-Wallis test shown in Table 5.4 indicate that the youngest students reported the highest self-efficacy levels in all four English skills. These results suggest that male students felt more self-efficacious than their female counterparts while self-efficacy decreased as the students age increased.

6.5.2.1 The students' self-efficacy as a function of gender

There is relatively very limited literature on how female and male learners of English as foreign language compare in terms of their self-efficacy for communicating in English. However, some literature indicates that male students usually feel more self-efficacious than female students particularly for completing tasks related to male dominated professions (Bandura, 1983; Schunk, 1995; Zimmerman, 1995). Similarly, female students have been found to demonstrate higher self-efficacy beliefs than their male classmates when tasks are thought to be in a female-domain (Onwuegbuzie et al., 2000; Pajares, 2002). In light of this literature, it may be argued that the self-efficacy differences observed between male and female learners in this study may be partly influenced by the learners' perceptions towards the use of English in daily activities. Culturally expected to act 'humbly' and stereotyped as less confident, quiet and shy (Warner, 2016), some Rwandan female students may feel less confident and not capable of proficiently

using English the speaking of which is still seen by some Rwandans as snobbish and a way of “showing off” (Kaneza, 2012). With English being spoken by only 1.9% of Rwandans in a predominantly monolingual society where 99.7% of the population speaks Kinyarwanda (MINECOFIN, 2005), Using it in daily life may be seen as unnecessary. On the other hand, male students who are culturally expected to be confident and more extrovert than female may see using English as being ‘cool’ (Kwibuka, 2013) and thus demonstrate more interest in English and feel more capable of using it proficiently than their female counterparts.

In addition, it seems that the observed differences may have also been influenced by other factors such as the students’ residence and the parents’ levels of education. In this study, the results show that students whose residence was in the rural areas had lower self-efficacy level for communicating in English as a whole than their colleagues from urban residences. Yet, 54% of the students who reported living in the rural areas were female. Also, the fact that fewer number of female students had parents with university education (43%) while the findings indicated that students who had parents with a university degree had higher self-efficacy (see Table 5.7), may be one of the factors behind the differences in self-efficacy levels between male and females. Although all these factors may seem to relate to the male-female gender differences in self-efficacy, the current literature suggest that learners’ cultural beliefs stands out to be the most consistent factor explaining the gender self-efficacy differences (e.g. Bandura, 1983; Onwuegbuzie et al., 2000; Pajares, 2002, 2003; Schunk, 1995; Zimmerman, 1995).

6.5.2.2 The age factor

Differences in students’ self-efficacy levels were also found to be the function of the students’ age. There is limited literature on the relationship between age and self-efficacy in EFL and no solid conclusion has been reached as to whether self-efficacy is influenced by age. In a study by Jenks (2004), 133 ESL students between the age of 11 and 20 from 3 Virginia public schools in the USA were surveyed to examine the effects of age, sex, and language proficiency on their self-efficacy. Results from Qui-Square test analysis found no effect relationship between the respondents’ age and their self-reported levels of self-efficacy. However, the researcher recognised that failure to identify any relationship could be due to the evenly

distributed age range and the broad range of the age scale that could have prevented any trend from emerging. In the current study, limited age range was used to maximize the accuracy of the measurement of trend. Participants were classified in 3 age groups of (1) 13 years and below, (2) between 14 and 15, and (3) 16 and above. Also, the Kruskal-Wallis test recommended for the analysis of group differences for more than two groups was used. The findings revealed that a large number of students aged 16 and above felt less self-efficacious than their younger classmates. The students aged 13 and below reported the highest self-efficacy levels of all the three groups (see Table 5.4).

The age factor revealed in this study can also be explained by some motivational theories that have demonstrated the influence of age on self-efficacy through learners' motivation and confidence to attain high achievement. It has been argued that while young students are process focused and can easily be motivated by rewards such as praise, older students are product oriented and poor achievement or rewards for success on easy tasks can easily lower their motivation and reduce their confidence (Barker and Graham, 1987; Meyer, as cited in Flammer, 1995; Pollard et al. 2000). Drawing on this literature, it seems that older students felt less self-efficacious for communicating in English based on their lower performance in English language assignments. In fact, the analysis of the relationship between the students' age and their self-reported score in the English subject revealed that some negative correlation existed between these two variable ($r = -.146, P=0.01$). In other words, the results showed that younger students aged 14 and below had higher scores in English as a subject than students aged 15 and above. This may have contributed in creating the feeling among older students with lower scores that they had no necessary abilities to successfully complete communication tasks in English.

It was also observed that there was an interaction between age and mode of school study. While the findings show that younger students reported the highest self-efficacy in all four English skills, the analysis of demographic data show that a large number of younger students were enrolled in boarding schools (84% of students aged 15 and below and 62% of students aged between 14 and 15). This seems to explain the reported higher self-efficacy levels of students in boarding schools. At the time of this investigation, students who were not

successful at the national ordinary level exam in Rwanda could either enrol in day schools (9YBE) or repeat the year to sit the exam for the second time. Many of the unsuccessful students opted to continue with day schools. Hence, day schools often had low performing students who were weak in English and who had not been able to continue in boarding schools. It is also important to note that some students in day schools were also returnees who had dropped school at some point for different reasons. All these factors seemed to explain why students at day schools often felt less self-efficacious for communicating in English.

6.5.3 Influence of school context

Following the differing characteristics of the surveyed schools, statistical tests were carried out to examine if the students' self-efficacy also differed as a function of school contexts.

Emphasis was put on two main school contextual factors of *learning mode* i.e. boarding school and day school and *school status* i.e. private, government aided or public schools. The Mann-Whitney and Kruskal-Wallis tests found statistically significant differences in the students' self-efficacy for using English in the four skills based on school contexts. It was found that the students from boarding schools reported higher self-efficacy levels than the students in day schools. The results also showed that the students from private schools reported the highest levels of self-efficacy.

Differences in self-efficacy between boarding and day school students could be influenced by a number of factors including the availability of resources, the students' demographic attributes as well as the students' educational background. It is important to note that unlike day schools many of which were introduced in 2009, Boarding schools in Rwanda are usually long established educational institutions. They often include private schools that are "administered and financed by private organisations or private individuals other than government" and government aided schools that are "administered and financed by the government with shares from private organizations or private individuals (Religious organizations, parents, etc.)" (MINEDUC, 2015b, p. 10). These schools usually enrol high performing students who have top scores in the national primary leaving examinations. This could suggest that the students' high performance was a factor that influenced their self-efficacy. In this study, the analysis of the relationship between the students' self-reported

scores and their self-efficacy indicated a strong predictive relationship between these two variables ($r=.52, p =.000$).

Therefore, it could be argued that, confident of having experienced high performance in English before, students in boarding schools felt highly self-efficacious while many of the students in day schools may be judging themselves on the basis of their previous failures on the national examination and hence feeling less capable of successfully completing communication tasks in English. This may seem to exemplify the mastery level source of self-efficacy (Bandura, 1997; Pajares, 1996; Zimmerman, 1995). By experiencing success, the students' self efficacy levels increase as a result and the opposite is true when they experience failure (Altschuler and Schmutz, 2006). Poor performance can decrease the students' confidence and creates feelings of incapacity that they cannot be successful in similar tasks. However, the students' self-efficacy reported in the current study does not seem to relate entirely to the students' past experience. The results show that students in most surveyed schools did not have enough opportunities to practice challenging and wide-ranging assessment activities in all four English language skills. Instead, they seem to base their self-report on their overall score in English summative assessment which was rather invariably focused on grammar.

Although many of the students at day schools were presumed to be low performing at the primary leaving examination, there were also students who had scored highly on this examination but decided to study at day schools because they could not afford the high cost of studying at boarding schools (Williams, 2015). However, the fact that they were not enrolled at boarding schools that were thought to be well staffed and having enough resources made some students at day schools feel weaker as explained by the Dean of studies at one day school in an ethnographic study by Williams (2015):

The challenges we face [as school administrators] here [at GSR] is that the students don't put good effort into their studies. They compare themselves with boardingschool students and feel that they are failures [...]. Their mindset is that they are here because they were not smart enough to get into a good school. 'Why am I even here?' they ask themselves. They feel that they are not clever... They think of themselves as students who will never do anything important for themselves. (Patrick, as cited in Williams, 2015, p. 14)

The students' feeling of weakness as stated in the quote above was also expressed by students in day schools when they were asked to say how they would compare themselves with students in boardingschools:

The government only prepares one national exam. We take the same exam done by the children of ministers who go to good schools [that have access to laboratories and other materials]. That is why the best students, those who perform well in national exams, are from cities. (Claude, as cited in Williams, 2015, p.9)

The lack of sufficient resources mentioned by Student Claude was also reported during teacher interviews. Teacher MR and SF indicated that materials such as exercise books were not readily available at their schools which appeared to be one of the challenges leading to the students' beliefs that they were not well trained enough to be at the same levels as other students in boarding schools. As the findings of this study tend to show, and as was implied in Williams' (2015) study, one of the effects that the scarcity of materials may have on the students is their feeling of inability and perceived low levels of self-efficacy.

In addition, differences in the students' self-efficacy based on school learning mode appeared to be also related to the education level of the students' parents. As demonstrated in Table 5.7, students whose parents had university education reported the highest self-efficacy levels. Yet, the descriptive analysis of the demographic information indicated that up to 66% of the 287 surveyed students from private schools were students whose parents had university education. However, these results do not mean therefore that the high self-efficacy levels reported by students were necessarily influenced by education levels of their parents as there is no substantive evidence from research so far that links the education level of parents to their children's self-efficacy in EFL. Instead, it seems that other factors relating to being highly educated such as the social economic status may be the most accurate influencers of the students' self-efficacy. In fact, some studies on learning motivation have found a positive relationship between EFL learners' motivation to use English and the social economic status (SES) of parents (e.g. Butler, 2015).

Butler's survey of 668 Chinese EFL learners from fourth, sixth and eighth grades found that students whose parents had higher SES provided their children with more opportunities to learn and use English in real life communication. This helped their children to develop high

self-determined motivation to learn and use English by taking part in conversations with native speakers, using English as they travelled abroad and seeking many more opportunities. As in Butler's study context, it is common that the highly educated parents in Rwanda are usually socially and economically well off and thus able to provide the same opportunities to their children. Given the importance of practice on the development of self-efficacy (Bandura, 1997, Pajares, 1996), it appears realistic to expect that students in Rwandan lower secondary schools whose parents had university education and opportunities to practice English outside the classroom reported higher self-efficacy...

The students' self-efficacy was also found to differ among students as a function of their school status where students from private schools reported the highest self-efficacy levels compared to the students in public and government aided schools. In this regard, the findings show that the gender of the teacher appears to be an important factor influencing the students' self-efficacy in private schools. While Table 5.9 shows that the students taught by male teachers reported higher self-efficacy levels than students taught by female teachers, data in Table 3.6 indicates that 86.7% of teachers in private schools were male, suggesting that the teachers' gender may be an influencing factor. However, this may also be mere coincidence and other factors may have led to these results. Very little research has been done on differences in the students self-efficacy based on the gender of teachers.

6.5.4 The influence of classroom-based assessment on students' self-efficacy

One of the major aims of the current study was to investigate the extent to which the teachers' classroom assessment practices, specifically their use of forms of assessment and their methods of providing feedback were related to the students' self-efficacy. This was particularly investigated in sub-research question 2.3 that was aimed at examining whether the teachers' use of assessment forms and feedback correlated with the students' self-efficacy for using English in speaking, listening, writing and reading. As presented in Section 5.3, The Spearman correlation analysis indicated that significant relationship existed between the students' self-efficacy and the teachers' use of some forms of assessment in the classroom. The fact that the teachers' frequent use of performance assessment tasks related to writing and speaking (as reported by the students) was identified as a possible factor that influenced the students' high

self-efficacy for speaking and writing underlines the potentials that performance assessment tasks offer to students when they are involved in performance activities that enable them to practice their skills.

As illustrated in some literature on self-efficacy development (e.g. Bandura, 1997; Schunk, 2003), productive skills are likely to be influenced by *performance* assessment because of potential opportunities to engage students in activities where they measure their abilities to execute tasks by practicing the use of English both orally and in written form. On the other hand, no correlation was found between the use of paper and pencil based forms and the students' self-efficacy for any of the English language skills. This was anticipated as the *paper and pencil* based assessment uses controlled forms of assessment such as *multiple choice* that do not offer opportunities for students to measure their ability to produce the learned language (through speaking and writing) or to understand written and spoken communication messages (through reading and listening).

While the analysis of data from students points to the relationship between the use of assessment forms with students' self-efficacy in writing and speaking (See Tables 5.12 and 5.14), it is important to recall that the analysis of data from teachers found positive correlation between the two variables only in the writing skill. In addition, no correlation was found between the use of assessment forms related to reading and listening and the students' self-efficacy for reading and listening in English respectively (see Section 5.3.1). These results do not seem to lend support to learner self-efficacy development theory specifying that students' self-efficacy is shaped by their experiences when they participate in classroom activities and observe others performing tasks (Bandura, 1997; Schunk, 2003). In fact, it was expected that teachers who reported high use of assessment forms in a given skill also had students with higher self-efficacy in such skill. This did not appear to be true as there was no correlation found between the use of assessment forms and students self-efficacy based on data from teachers.

Similarly, the analysis of data from teachers did not reveal any relationship between the teachers' reported methods of providing feedback and their students' self-efficacy. As was

expected for assessment forms, it was also expected that students whose teachers reported high use of diversified feedback and less reliance on marks would report the highest self-efficacy. However, the results did not show any relationship between the teachers' use of feedback methods and their students' self-efficacy based on data from teachers. On the other hand, while these results suggest that the nature and type of feedback did not seem to affect the students' self-efficacy, data from students indicate that frequency of teachers' feedback seem to affect the students' self-efficacy. As indicated in Table 5.17, the students who believed that they received frequent feedback from teachers also reported high self-efficacy.

These inconsistencies seem to be a result of effects from 'social desirability bias' on the side of teachers. As was highlighted in section 7.3, some teachers appeared to overstate the frequency of their use of assessment forms and feedback apparently due to 'social desirability bias'. Although they may not be able to use them in their classroom for different reasons, teachers may be aware of the important role that the use of some feedback and assessment forms plays in promoting learning. Teachers may have wanted to show that such forms of assessment and methods of providing feedback were used in their classrooms in order to meet the expectations of the researcher. In addition to possible effect of 'social desirability bias' on the side of teachers, the lack of relationship between feedback and the students' self-efficacy may also be due to the nature of the provided feedback, and the value that the students attached to it. While some argue that teachers' feedback enhances the students' judgement of their own abilities and leads to the development of self-efficacy (Ross et al. 2002), it appears that the lack of formative feedback prevented students from being able to accurately judge their abilities based on practical facts. The students seemed to be interested in numerical feedback in the form of marks which they considered as a good indicator of their abilities.

Overall, it can be concluded that the students' judgement of their self-efficacy for communicating in English was influenced by several factors. These include factors related to the students' educational background and demographic features as well as some school contextual factors. The latter were also found to affect the teachers' assessment practices in the classroom. The findings also point to the limited assessment literacy of teachers which, in conjunction with some school related factors, influenced their classroom assessment decisions

which in turn had effects on students' beliefs (self-efficacy). From a more general perspective, the findings of the current study seem to suggest that the students' self-efficacy for speaking, listening, writing and reading in English language was affected by not only what happened but also what did not happen in relation to classroom assessment. On the one hand, the prevalent use of paper and pencil based assessment and the related sense of achievement experienced by students appeared to lead to their beliefs that they could successfully execute real life communication tasks in English. This was contrary to the commonly held belief in the country that students demonstrated poor performance in English communication (Osae, 2015; Tabaro, 2012).

On the other hand, the limited use of wide ranging assessment tasks to create opportunities for students to practice the language real-life like communication seem to have led to the students' misjudgement of their self-efficacy levels. As Bandura (1983) stresses, performance opportunities and sufficient performance information are some of key factors that can help avoid inconsistency between self-efficacy and real performance. He argues that "discrepancies may arise because of misjudgment of task requirements, unforeseen situational constraints on action, inadequate tools and resources for optimal execution of skills, deficient performance information so that self-percepts of efficacy are not translated to appropriate magnitudes of effort" (Bandura, 1983, p. 251). Learning English in a predominantly monolingual society such as Rwanda, many Rwandan students use the classroom as their only place for practicing English communication as there are very limited opportunities outside the classroom where they can use English. Students are also likely to get information on their abilities to perform communicative tasks in English from their classroom assessment information specifically during feedback from both teachers and peers. It sounds realistic therefore, that the insufficiency of this information or its irrelevance might lead to the misjudgement of the students' own abilities.

6.6 Summary of the chapter

This chapter discussed further the results presented in chapter four and chapter five. It argues that the prevalent use of paper and pencil assessment forms and the marks-based feedback in the lower level of the secondary school in Rwanda reflects a common practice in high stakes

testing contexts. The current chapter argues that the test washback effects appear to explain why teachers seem to use more controlled assessment forms than performance based forms. However, the classroom contexts such as large class size and the contextual values are also some of other important factors that may explain the dominant use of controlled forms of assessment. It is also argued in the current chapter that the inconsistency identified between the teachers' self-reported assessment practices and their real practices during classroom assessment may be a result of the teachers' apparent limited assessment literacy as has been found in other previous studies in similar EFL contexts (e.g Mekonnen, 2014; Qassim, 2008). One of the important findings in the current study was the lowest satisfaction of the students in the assessment transparency yet they reported high performance on assessment assignments. This may demonstrate a gap between understanding and performance and may be a reflection of the consequence of a product oriented assessment. It also seems to underline the importance of the learner consultation and involvement in the assessment process to reinforce understanding and points to the need for calibration practices (Pajares, 1996; Phakiti, 2006) to train learners on how to accurately judge their abilities. On the relationship between assessment practices and the learners' self-efficacy, the results in chapter five show that the influence of the external factors on the students' self-efficacy seems to outweigh the influence of the teachers' classroom assessment practices in the context of Rwanda. However, the observed statistically significant relationship, though marginal, between the use of assessment practices and the students' self-efficacy in speaking and writing underlines the important role that assessment can play in increasing the students' self-efficacy.

Chapter 7: Conclusion

7.0 Introduction

Chapter one and two presented the contextual background and the conceptual framework of the current study. Chapter three highlighted the methodological approaches that were adopted for the process of data collection and data analysis. The results of the study were analysed and presented in Chapter four and five while findings are discussed further in chapter 6. In the current chapter, a brief recap of the major findings is provided and more attention is drawn to the implications of these findings as well as the main contribution of the study as a whole. The chapter also highlights the major limitations of the study and suggests areas for further research.

7.1 Summary of the current study findings

The purpose of the current study was to explore the classroom assessment practices of teachers of English at the lower level of secondary schools in Rwanda and to investigate the extent to which those practices were related to the students' self-efficacy in the four English language skills. Data was obtained by means of teachers' self-report, classroom observation, teacher interviews and students focus groups. On research question one that investigated the teachers' classroom assessment practices and the students' perceptions, the descriptive analysis of the quantitative and qualitative data indicated that the teachers' classroom assessment was characterised by a predominant use of *paper and pencil* forms of assessment (see Section 4.1). It was also found that the teachers had little knowledge about the use of rubrics and therefore hardly used them for their classroom assessment. The teachers' methods of providing feedback also appeared to be dominated by the use of marks and verbal teacher feedback usually addressed to the whole class. These results are consistent with the findings of studies in EFL contexts that have demonstrated a common tendency among EFL teachers to use more *paper and pencil* based assessment than *performance* based assessment (Cheng et al., 2004; Frey and Schmitt, 2010; Tante, 2010; Tzagari and Pavlou, 2009). The current results were also expected given some contextual factors that are perceived as promoting the use of *paper and pencil*

assessment forms. Some of these factors were cited by teachers during their one-to-one interviews with the researcher. These included the lack of assessment resources, the teachers' limited assessment literacy and the effects of the top-down assessment system where high-stakes national tests were found to affect the teachers' classroom assessment practices (see Section 6.2.2).

On the other hand, the students reported rather positive perceptions of the teachers' classroom assessment practices with respect to the authenticity of assessment tasks, congruency with the learning content, assessment frequency, fair marking and teacher feedback. The overall high level of students' satisfaction of the classroom assessment was less expected. The students' perceptions seem incongruent with the reported classroom assessment practices presented as being highly grammar focused, favouring the use of *paper and pencil* assessment and using a very limited range of assessment tasks with teachers' feedback invariably based on marks. One possible explanation of this discrepancy in this study context may be that the students did not have the necessary ability to appraise their classroom assessment due to their limited assessment literacy. It appeared that the students were only aware of and familiar with the commonly used paper and pencil assessment with more importance given to marks than the real demonstration of improved competence. The lack of awareness of the wide ranging assessment forms and the potentials that they can offer them might have influenced the students' responses.

On research question two that investigated the students' self-efficacy and how it was affected by their teachers' assessment practices, the results showed that overall the students reported high levels of self-efficacy in all four English language skills especially in reading and speaking. The significantly high levels of self-efficacy reported by the students in all four skills were highly unexpected. Drawing on my personal experience both as a learner and as a teacher of English at the lower secondary school level and based on the available published reports and studies that often point up the perceived Rwandan students' limited ability to use English (Basheija, 2014; Osae, 2015; Tabaro, 2012), it was not anticipated that the students would report such high self-efficacy in using English especially in the productive skills (i.e. speaking and writing). In addition, the results showed that the students did not get enough opportunities to perform tasks related to all four language skills and therefore might not have full

understanding of their performance abilities. In this regard, the high self-efficacy levels reported by the students were interpreted as being based on the teachers' provided information from assessment which usually consisted of marks and a few generic comments. The comments from the majority of students during focus group interviews suggested that the students used the marks from classroom assignments as their 'subjective standards of success' (Ross et al., 2002, p.83), and as an indicator of their general ability to successfully perform tasks in all the four skills. It may be important to recall that the students reported high grades in English language assessment in general which appeared to enhance their beliefs about their self-efficacy in the four English language skills. This seems to confirm that self-efficacy can be overestimated when there are no opportunities to demonstrate the real levels of one's ability (Pajares, 1996). The students may have overestimated their self-efficacy due to the limited awareness of the required abilities associated with real performance of given tasks.

The correlation analysis of data from the students' survey indicated that the differences in the students' self-efficacy were based on various factors including mainly the school related and the students demographic factors (see Sections 6.5.2 and 6.5.3). The school learning mode (boarding vs day School) was found to be the most important school related factor that appeared to influence the students' self-efficacy. Differences based on school learning mode essentially resulted from other issues associated with factors such as resources availability. It was often said that teaching and learning resources were more readily available at boarding schools than at day schools. However, it is also important to recall that Boarding schools enrol the best students who have the highest grades on the National examinations. As explained earlier, high marks and grades were considered by many students as a source of information about their overall ability to perform communicative tasks in English. This may have significantly affected the students own judgement of their self-efficacy in English language skills.

Concerning the relationship between the assessment practices and the students' self-efficacy, the analysis of data from the students suggest that a degree of relationship existed between the use of performance assessment forms pertaining to speaking and writing and the students' self efficacy in speaking and writing respectively. On the other hand, the results from the analysis of teachers' self-reported data showed that a relationship existed between

assessment practices and the students' self-efficacy only for writing. No relationship was found between the students' self-efficacy for listening with the use of listening related assessment forms. This was not surprising as listening was rated by both teachers and students as the least assessed skills. Nonetheless, the low correlational coefficients (below $r=.40$) indicated a weak level of associations between assessment practices and students' self-efficacy apparently due to unbalanced use of assessment forms.

7.2 Implications of the study

Drawing on the study findings summarised above, a number of educational implications are suggested. First, it was clear from the findings that some teachers at the lower level of secondary schools in Rwanda had divergent interpretations of assessment. This was discussed in earlier chapters as 'limited assessment literacy' (Section 6.2.1) on the side of teachers. One implication of this might be the introduction of training programmes that would have as a primary purpose to provide professional development support to teachers in order to advance their assessment literacy. This may call for training approaches aimed at developing "knowledge, skills and abilities required to design, develop, maintain or evaluate, large-scale standardized and/or classroom based tests..." (Fulcher, 2012, p. 135). It is essential that teachers are provided with clear guidance for assessment practices that reflect the processes, principles and concepts of effective classroom assessment. The professional support can be provided in different ways including enabling teachers to access online courses that provide teachers with the opportunity to explore the latest approaches and skills in relation to language assessment (e.g the online Teachers' Assessment Literacy Enhancement [TALE] project). Online courses can be convenient for many teachers who may need to complete quality training at their convenient time while also fulfilling their teaching responsibilities on a regular basis.

Second, in addition to providing in-service training to raise the assessment literacy level of teachers, it is also vital to provide them with the necessary resources including mainly the assessment materials that cater for their needs in their everyday assessment practice (Fulcher, 2012). The results of the current study indicated that there was a need for the supply of enough teaching materials adapted to the level of the students and designed based on their language needs. It was reported during the survey that most schools lacked sufficient and appropriate

materials meaning that teachers were compelled to look for other sources of assessment. The provision of teachers' guides on classroom assessment would encourage teachers to use more appropriate assessment methods and help them enhance their assessment skills as they assess their students. The findings indicated that teachers chose to stick with controlled assessment partly because these were more easily accessible than performance based forms. However, it was also observed that some teachers chose to use the assessment methods similar to those used in the national examinations. This may call for a change in the format of the national examinations that need to as comprehensive as possible in relation to language skills. As has been discussed in earlier sections (e.g see section 2.1.3.2), the high stakes tests can have significant influence on the classroom instructional practices including the teachers' methods of assessment. It can be argued therefore, that the change in the format and content of the national examinations coupled with the teacher training and supply of materials would bring significant improvement in how teachers of English teach and assess their students. The teachers' use of diversified assessment materials would have a positive impact on the learners' assessment literacy as it would increase their awareness of a wide range of assessment forms and give them the opportunity to practice the language through rich and diversified assessment activities. In this study, it was observed that, having been exposed to a very limited range of assessment, the students had very limited knowledge of the possible alternatives for their classroom assessment. The students' full engagement in the assessment process through selection of assessment activities, use of peer assessment and familiarity with use of assessment rubrics can also contribute in raising the students' assessment literacy levels (Smith et al. 2013).

Third, it was found that despite the well known lack of confidence in Rwandan learners to communicate in English, the students reported relatively high levels of self-efficacy in all four English language skills. This was found to be a result of the high value that the students attached to marks and grades upon which they based their confidence for being able to successfully complete tasks in English language. The students' high reliance on marks and grades for information on their abilities is also related to the paucity of diversified assessment opportunities. The students did not experience assessment tasks that are comprehensive and challenging enough to provide them with high quality information on their real abilities to communicate in English which led to the misjudgement of their self-efficacy. This highlights a

need for a course of actions to transform the assessment practices of teachers so as to put more emphasis on assessment practices that provide sound information to the students on their real abilities in relation to using English in all the four skills. As Zimbicki, (2007) emphasizes, teachers need to create assessment opportunities that are not only based on traditional paper and pencil assessment but also engage students in alternative assessment:

Engaging students in meaningful learning experiences and using alternative assessment throughout the learning process may have a significant impact on motivation and self-efficacy [...] Teachers need to de-emphasize performance on traditional paper-and-pencil tests, and instead use authentic forms of assessment making learning more relevant and meaningful to students [...] Students need to participate in assessment activities that provide them with opportunities for success for a better realization of their intellectual development and self-efficacy. (p. 7).

In addition to engaging students in diversified assessment activities, teachers also need to ensure that the students get quality feedback that contains enough information on their performance in order to avoid misjudgment of self-efficacy. In the absence of other sources of information on their performance, the students in this study showed that they considered the marks and grades usually obtained from grammar based assessment as convincing evidence of their overall ability to use English effectively. Teachers should give quality feedback that provides learners with information on their actual progress vis-à-vis the learning goals and clarifies the required steps to achieve the target goals (Hattie and Timperly, 2007). Quality feedback can help raise the accuracy of the students' self-perceptions, also called 'students' calibration', by helping them to understand what they know and what they do not know thereby avoiding overestimation of their self-efficacy (Pajares, 1996; Phakiti, 2006).

7.3 Limitations of the study

It is important to acknowledge that the findings of the current study are subject to some limitations most of which were associated with the entire process of data collection. The findings presented in this study have been reached after a research process that used different methods and instruments for collecting and analysing both quantitative and qualitative data. However, although the use of mixed methods was advantageous in that it helped to investigate the research questions in some depth and allowed for triangulation of data, it appeared that the size of the collected qualitative data was smaller than the desired size. The classroom

observation, the teachers' interviews and the students' focus groups were initially intended to supplement the quantitative data by gathering a sufficient amount of realistic information on the classroom assessment practices. However, it is important to acknowledge that some unpredicted setbacks led to a restricted size of data being gathered hence calling for cautious interpretation of the current results.

One of the major setbacks was the lengthy and exigent process of obtaining the research ethics clearance from the local research ethics committee and the reluctance of some schools and teachers to participate in the study partly due to their busy schedules. A limited number of teachers agreed to be interviewed and fewer lessons were observed for shorter periods of time than was initially planned. Short time of observation could mean that some features of the teachers' classroom assessment practices might not be captured as desired. I believe that rich and comprehensive data could have been obtained if more teachers and students from a wide range of schools were reached, interviewed and observed for a sufficient amount of time. The process of data collection in this study was time bound because it had to be completed within less than three month period which was the normal length of one school term. This was a very short time considering that it was within the same period that the lengthy process of obtaining the consent of the students' parents needed to be completed.

The generalisability of the current results is also subject to certain limitations. As the study used convenience sampling techniques, the research sample was not entirely representative of all teachers and students of the lower secondary schools in Rwanda. Therefore, the current findings may not be applicable to all classroom assessment contexts of the lower level of secondary schools. This study was conducted in seven of thirty districts of Rwanda representing only 23% of the country's geographic area. Therefore, the results may not necessarily be generalised to other education districts of the country. Nevertheless, these findings may be of interest and relevance to many districts and schools that share the same characteristics with the surveyed schools. The scope of the current study was also limited in terms of the focus of investigation. The research emphasis was put on the teachers' and students' perspectives about the classroom assessment practices in relation to the forms of assessment and methods of feedback. It was not specifically designed to explore the classroom assessment from the perspective of the material developers or policy makers at the top level. As

a result, the current study did not particularly seek to obtain or use data on textbooks or assessment policy at the national level. The data analysis was only limited to information reported by the respondents and data from observations. It might be possible therefore that the study missed some details that would have provided supplementary information for a broader understanding of the context.

It is also important to bear in mind that there might have been possible bias in the answers of the respondents particularly the teachers during their questionnaire self-report. They may have wanted to evade answers that they thought would be perceived negatively by the researcher. Given the discrepancies between data obtained from the students and classroom observation on the one hand, and the teachers' self-reported data on the other hand regarding the use of performance assessment in the classroom, it is plausible to suggest that the teachers' self-report may have been subject to the 'social desirability bias' (See Section 3.3.1 and 3.8). Knowing that I worked in higher education, some teachers might have considered me as being more trained than they were and therefore did not want to reveal their actual classroom assessment practices over concerns that I might judge their competence. This is commonly referred to as 'power relationship' between the researcher and participants that can sometimes affect data when it is highly imbalanced (See section 3.8). This raised issues of reliability of some information provided by the teachers as they seemed to suggest that they frequently used *performance* assessment for classroom assessment while data from the students and classroom observations suggested that *paper and pencil* assessment were predominantly used. Although the qualitative data from classroom observations and interviews was used to validate data from the questionnaire surveys, the quantitative data remained central to the current study especially with regard to the teachers' classroom assessment practices and the students' self-efficacy. The accuracy of some reported data may thus not be fully validated and calls for careful interpretation of the results.

7.4 Contribution of the study

Notwithstanding the limitations outlined above, the current study makes noteworthy contributions to the current literature on the classroom assessment practices of EFL teachers and extends our understanding of the students' self-efficacy in EFL context. The current

findings are of a particular importance to the Rwandan context as the study appears to be the first empirical research to report on the teachers' classroom assessment and learners' self-efficacy in Rwanda. The few studies that have been conducted on assessment in Rwanda have almost entirely focused on higher education overlooking the secondary school in general and the lower level of secondary education in particular. With English language teaching and learning still in its infancy in this country that has known French and Kinyarwanda as the only official languages and media of instruction for decades, the findings of this study offer some major insights into the English language instruction and sheds some light on the current implementation of assessment in Rwandan English language classroom. The current findings also make very important contribution by producing more literature on English language instruction and assessment in similar contexts of English as a foreign language. The study may also serve as a base for further studies on the students' self-efficacy which seems to be vital to the Rwandan context as a monolingual society. Rwandan learners' of English need to develop their self-efficacy in order to be able to persist and succeed in their English language learning.

As was highlighted in the introductory chapter of the study (section 1.5), this study was carried out following the researcher's observation that there was a lack of literature on self-efficacy and classroom assessment practices in EFL contexts especially for the secondary schools level. In addition, while assessment has been viewed as an essential component of instruction and self-efficacy considered as a predictor of performance, very little has been done to explore the possible relationship between the two (see section 1.5). This may be partly due to the fact that self-efficacy as a research topic is quite new in EFL. The current study produces more literature in this respect and furthers our understanding of this relationship. The findings of the current study also highlight the relationship between the students' perceptions of classroom assessment and the students' self-efficacy. The findings reimphasize the importance of Bandura's (1997) self-efficacy theory particularly on the role that the enactive mastery experience plays to the development of learners' self-efficacy. The fact that the students in current study based their self-efficacy judgement on their performance (grades and marks) highlights the crucial influence that the experience of what individuals judge as success has on their self-perceived abilities. In addition, this study provides additional evidence that external factors can also be influential to the development of the learner self-efficacy and the role of learners' assessment literacy in the understanding and judgement of their own self-efficacy.

The current findings also draw attention to the fact that the lack of effective feedback can lead to the students' misjudgement of their abilities to perform communication tasks.

In addition, despite the methodological limitations highlighted in Section 7.4, the use of mixed methods in the current study helped to generate some rich and diversified sets of data and allowed for the control of the weaknesses inherent in each method when used alone (Cohen et al., 2007; Dornyei, 2007). As was highlighted earlier in section 1.5, most related studies have relied on one research method to investigate the classroom assessment practices and the students' self-efficacy and this can raise the issue of data reliability. The current study has reemphasized the importance of mixing both quantitative and qualitative research for in depth analysis of information through data triangulation. One of the strengths of the current study was that data was gathered in four different ways (questionnaire, classroom observation, interviews and focus groups) which enabled the researcher to authenticate different bits of the respondents' information which increased the researcher's confidence in drawing relevant conclusions.

7.5 Suggestions for further research

The current research has attempted to answer two important questions on EFL classroom assessment and EFL learners' self-efficacy particularly in the context of the lower level of secondary schools in Rwanda. However, more questions have been raised as a result and further studies are needed to provide answers to those questions and issues. In addition, the significance of the limitations encountered during the current research process suggests that further investigations are warranted in order to obtain robust and more reliable evidence. For example, on the methodological level, it is my personal belief that a future study investigating the classroom assessment practices using qualitative data as a primary method for data collection would yield robust and more reliable evidence. It has been explained earlier that the current study relied heavily on quantitative data that was collected through the questionnaire survey where in some cases the accuracy of the respondents' information could not be sufficiently verified. A qualitative approach should attempt an investigation of the classroom assessment where more classroom observation and other more suitable methods of data collection and analysis could be used to get sufficient data on the actual assessment practices of

EFL teachers. For example, the ‘document analysis’ is one of the important methods that could be used to gather more accurate and reliable data on the teachers assessment practices in their classrooms.

Another possible methodological variation for the future research would be to carry out a study similar to this one using experimental methods to measure the impact that the use of different forms of assessment can have on learners’ self-efficacy. Instead of asking the respondents to provide a self-report on what they do in the classroom in relation with assessment, I suggest that an experimental study should use predesigned tasks in the classroom over a period of time. This experimental intervention would use well designed *performance* based and *paper and pencil* based assessment aligned with the measures of the students’ self-efficacy. Appropriate alignment between measures of self-efficacy and measures of performance is recommended in self-efficacy studies as a way of enhancing the reliability of results (Bandura, 1986; Moriarty, 2014). In the current study, a generic inventory of assessment forms were presented to the teachers and it was presumed that teachers used some of the suggested forms to assess their students. Similarly, the statements used to measure the students’ self-efficacy were selected from topics outlined in the national curriculum for English language at the lower level and it was expected that teachers had assessed their students on the same topics prior to the research. These would be better controlled in the experimental intervention to ensure accurate alignment of assessment forms and measures of the students’ self-efficacy before data could be collected.

As explained in Section 7.4, the scope of the current study was also limited both in terms of subject matter and geographical areas of the survey. The investigation was mainly focused on the practice of assessment as it happened in the classroom and focused on a limited area of Rwanda which in turn limits the generalisability of the results. One way of increasing the generalisability of findings would be to expand the geographical scope of the study and to include a maximum number of schools that best represent the features under investigation. This is often referred to as ‘studying the typical approach’ where priority is given to gathering valid and reliable data over convenience and ease of access to data (Schofield, 2000). It is important therefore that future studies identify typical schools contexts across Rwanda where teachers from different background can be given the opportunity to provide information on their

classroom assessment practices and the students to report on their assessment experiences and self-efficacy. Moreover, future studies may seek information on the assessment policy to help get a broad understanding of the context in which teachers work. A progression of the current study could therefore investigate the language assessment policy in Rwanda and explore how this policy is implemented by teachers at the lower level of secondary schools. The study may also look at the role of school administrators to whom teachers usually report the assessment results. The participation of the school administrators and respondents at the policy making level would help identify the political and educational factors that influence the teachers practices in the classroom.

More research is also needed to better understand the role that self-efficacy plays in motivating EFL learners to use English language in the four skills. Research should examine for example the extent to which the learners' self-efficacy for speaking affects the learners' actual speaking of English in their everyday communication. According to the findings of this study, it would be expected that higher levels of self-efficacy as reported by the learners would mean that they also exhibited strong enthusiasm to speak English. However, it was often reported that Rwandan students were often reluctant to use English whether in school environment or outside the school (Basheija, 2014; Osae, 2015; Tabaro, 2012). Drawing on the current findings that provide some insights on the learners' self-efficacy, a separate study is needed to explore the role that the learners' self-efficacy plays in the eventual use of English.

7.6 Personal reflection

This section presents my personal reflection on the entire journey that led to the current thesis. It revisits the process through which this research project was carried out and describes the important aspects of the journey characterized by tense moments of anxiety but also marked with feelings of achievement and satisfaction. From the time I took the decision to undertake the PhD degree programme in language education, I was aware that my new academic undertaking was going to be challenging but also rewarding and highly transformative both for me and for my young family. However, the successful completion of the previous postgraduate degree programme two years earlier and the high interest that I had developed in educational research all gave me the courage to challenge myself with a new academic endeavour. I also

understood that my personal circumstances were going to be different and would require strong commitment and undivided attention to be able to arrive at a successful end of the programme.

Every stage of the study was a learning experience starting with the decision itself of leaving my job and coworkers once again for another three year scholastic project. Nonetheless, I was well aware that it was an unmissable opportunity in order to have my dreams come true. After my master's studies in 2012, my goal was to embark on a doctoral study programme before I got busy with other life ambitions. I realised my dream when I received the offer to study at the University of York. I was then confident that my research project would kick start smoothly and move through stages in a minimum of time to finish exactly in three years if not slightly earlier. However, when I arrived at the university and got access to the latest and state of the art literature about my research area, I realised that my then suggested topic of study was too broad and could not possibly be undertaken as it was. It became clear that the preliminary literature review that I had done when I was writing my proposal was not extensive enough and did not provide me with the full portrait of research in the area. My challenges had started already as I was compelled to change or modify the topic. Although this was unexpected, I considered it as an earlier warning signal of the critical role that extensive reading has to play regardless of the stage of the project. It appeared that, as a novice researcher, I did not have a full grasp of the high extent to which critical reading of literature is crucial in earlier stages of the study and during the entire research process as a whole.

Each experience at every step informed the next stages and this was particularly the case when it came to redesigning data collection instruments and choosing appropriate methods of data analysis. My initial plan was to use quantitative methods only but as I read the literature I became more aware of the shortcomings of relying on one method and realised the benefits of utilising various instruments for data triangulation. I therefore decided to use mixed methods in an effort to gather as much valid and relevant data as possible. The process of collecting data was both informative and satisfying as it was my first time to do research at such a large scale. As a former secondary school teacher myself, I was familiar with the secondary school system in the country and I knew where most schools were located. I was aware of some of their normal routines which helped me to plan school visits. However, the whole process was not

without problems. One particular challenge that I faced surfaced when I applied for the local ethics clearance before I could start collecting data in Rwanda. I had to undergo unnecessary delays while the institutional review board was in consultation with other boards within the university to decide who had legal competence to issue the clearance. This was to an extent disappointing although I knew that it was one of those problems that could occur along the way. I am however grateful that the permission was later granted and data was collected according to plan.

In short, the three years of my research journey has been transformative not only on the research level but also on my personal growth as young professional who ought to face challenges and seek solutions to problems as they arise. In addition to building a wide and solid knowledge foundation in my research area, I have strengthened understanding of the importance of having a dream, persistence and seeking collaboration with others. To me the research process that resulted in the current thesis was a good opportunity to exercise my decision-making skills and I was aware that any decision made at any stage of the research would have significant impact. It was very important to remember that every decision that I made needed to be given serious consideration, starting from the time of deciding on the research topic to the moment of choosing the date of thesis submission. I am confident that the wide knowledge and skills that I got as a result of a successful completion of this project will guide my future undertakings in pursuit of commendable achievement.

Appendices

Appendix 1A: Questionnaire for students (English version).....	248
Appendix 1B: Questionnaire for students (Kinyarwanda version).....	250
Appendix 2A: Student focus group schedule (English version).....	253
Appendix 2B: Students focus group schedule (Kinyarwanda version).....	254
Appendix 3A: Students' assent form (English version).....	255
Appendix 3B: Students' assent form (Kinyarwanda version).....	256
Appendix 4A. Classroom Observation scheme: Assessment forms used in the classroom...	257
Appendix 4B. Classroom Observation scheme: Classroom oral interactions.....	258
Appendix 5: Teacher questionnaire.....	259
Appendix 6A: Consent form for teachers (English version).....	264
Appendix 6B: Consent form for teachers (Kinyarwanda version).....	265
Appendix 7A: Consent form for Head teachers (English Version).....	268
Appendix 7B: Consent form for Headteachers (Kinyarwanda version).....	270
Appendix 8A: Consent form for parents (English version).....	272
Appendix 8B: Consent form for parents (Kinyarwanda version).....	274
Appendix 9: Sample teacher interview transcript (Interview with Teacher JS).....	276
Appendix 10: Sample transcript from the student focus group.....	279
Appendix 11: sample qualitative data analysis: codes and categories.....	283
Appendix 12: Instruments and data collection methods used for each research question....	286
Appendix 13: Results from Spearman test of correlation between the teachers' use of performance assessment forms related to reading and the students' self-efficacy for reading in English.....	287
Appendix 14: Results from Spearman test of correlation between the teachers' use of performance assessment forms related to listening and the students' self-efficacy for listening in English.....	288

Appendix 1A: Questionnaire for students (English version)

Please answer as per instructions for every question.

PART I:

Please tick the appropriate box or write your answer in the box

1. Gender: Female Male

2. Age:

3. Average mark in English language for the last term (e.g. 15/20 or 70%)

4. Which category best describes the highest level of education of your parents? Please tick the in box

Primary education ; Secondary education , University degree

5. Which of the following best describes your home area?

a) Country side? , b) city/town

PART II.

6. Does the content of classroom assessment match the content of instruction? (Underline the answer)

a) Strongly agree b) Agree C) Disagree d) Strongly Disagree

7. How do you think English will help you in the future?

8. Do you think the current classroom assessment help you in that respect? (Please underline the answer)

a) Strongly agree b) Agree C) Disagree d) Strongly Disagree

9. Please tick the box corresponding to your chosen answer.

	Statement	Strongly Agree	Agree	Disagree	Strongly Disagree
a	I take part in deciding what needs to be on my assessment activity and how to do it				
b	It is always clear what the teacher wants me to do on assessment activity				
c	The teachers uses different kinds of assessment activities (example: written tests, oral presentations, etc)				
d	I always get feedback from my teacher after assessment				
e	I think the teacher always grades my work fairly				
f	Overall, the teacher gives us enough assessment activities				

10. Please tick in the box to show how frequently these forms of assessments are used in your classroom

		Usually	Sometimes	Rarely	Never
1	Written assessment E.g: fill in questions, underlining right or wrong answer, multiple choice questions, etc)				
2	Reading assessment E.g. reading comprehension test				
3	Listening assessment. E.g. Listening to audiotapes				
4	Writing assessment. E.g paragraph writing				
5	Role play E.g. Performing a written dialogue				
6	Speaking test. E.g. Oral presentations, debates				

PART III

11. On a scale from 0 (0 meaning *no chance at all*) to 100 (100 meaning *completely certain*), **how sure are you that you can** perform each of the tasks below in English easily? Please choose the number that best reflects your belief.

	Tasks	Level of certainty
1	I can orally describe members of my family in English	
2	I can talk about the consequences of lack of hygiene	
3	I can give a talk on how I spent the last holidays	
4	I can orally explain the importance of water	
5	I can orally discuss the advantages and disadvantages of living in a city or in a country side	
6	I can write a letter to a friend explaining how I spent the last holidays	
7	I can write a paragraph about the consequences of lack of hygiene	
8	I can write a story about my first day at the secondary school	
9	I can write a paragraph describing what the head teacher said in a meeting	
10	I can write a summary of what is said in any meeting held in English	
11	I can read and understand a text on the responsibilities of parents in a family	
12	I can read and understand the biography of my favourite singer	
13	I can read and understand a text about education system in Rwanda	
14	I can read and understand a text comparing describing the hills of Rwanda	
15	I can read and understand a paragraph describing the importance of water	
16	I can listen to and understand an audio about the role of adults in the country	
17	I can listen to and understand what is said in the news on Radio Rwanda	
18	I can listen and understand what is said in a radio broadcast about education	
19	I can listen to and understand a audio discussion describing the hills of Rwanda	
20	I can listen to and understand radio broadcast in English on Radio Rwanda	

Appendix 1B: Questionnaire for students (Kinyarwanda version)

Subiza ukurikije amabwiriza yo kuri buri Kibazo. Si ngombwa kwandika izina ryawe kuri uru rupapuro.

IGICE CYA MBERE

1. Erekanako umwirondoro wawe ushyira akamenyetso kugisubizo kiri cyo:

a) Gabo (male) , b) Gore (Female)

2. Imyaka ufite (age)

3. Wagize amanota angaha mu isomo ry' icyongereza mugihembwe gishize (urugero 70%)

4. Umubyeyi wawe cyangwa uwo mubana murugo wize amashuri menshi yarangirije mukihe kiciro?

a) Amashuri abanza ; b) Ayisumbuye , c) Kaminuza

5. Ni ikihe muribi gisobanura aho utuye iyo utari ku ishuri? Shyira akamenyetso kugisubizo

a) icyaro b) Umugi

IGICE CYA KABIRI

Subiza ibibazo bikurikira

6. Ubona ibyo mwibandaho mwiga mucyongereza ari nabyo mubazwaho kenshi? (Ca akarongo kugisubizo)

4) Yego cyane 3) Yego 2) Oya 1) Oya rwose

7. Ni akahe kamaro wifuza ko kumenya icyongereza byazakumarira mugihe kizaza?

8. Ubona uburyo ubazwamo buzagufasha kuzabigeraho? (Ca akarongo kugisubizo)

4) Yego cyane 3) Yego 2) Oya 1) Oya rwose

9. Andika akamenyetso mukazu kajyanye nigisubizo cyawe.

		4. Yego Cyane	3. Yego	2. Oya	1. Oya Rwose
a	Njya ngira uruhare muguhitamo ibyo tubazwaho n'uburyo bwo kubazwamo				
b	Burigihe mwibazwa mba numva neza icyo mwarimu ashaka ko dukora				
c	Mwarimu agenda ahindura uburyo bwo kutubaza				
d	Burigihe mwarimu w'icyongereza ampa ibisobanuro by'uko nakoze mwibazwa				
e	Muri rusange nyurwa nuburyo mwarimu w' icyongereza akosora				
f	Muri rusange duhabwa ibazwa kuburyo buhagije				

10. Erekena inshuro mwishuri ryawe mubazwa hakoreshejwe ubu buryo. Shyira ✓ mukazu k'igisubizo

		4.Akenshi	3. Rimwe na rimwe	2. Gake cyane	1.Nta na rimwe
1	Ibazwa ryo kwandika ibisubizo (Urugero: teste yanditse irimo ibibazo byo kuzuza ahatuzuye, guca akarongo kugisubizo, guhitamo igisubizo kiricyo n'ibindi)				
2	Ibazwa ryo gusoma (reading) . Urugero: teste yo gusoma no kumva umwandiko				
3	Ibazwa ryo kumva ibivugwa (Listening) . Urugero: teste yo kumva ibivugwa kuri radio cyangwa kuri tape (cassette)				
4	Ibazwa ryo kwandika (writing) . Urugero: teste yo kwandika ibaruwa, inkuru cyangwa inshamake				
5	Ibazwa ryo gukina (role play) . Urugero: Teste yo gukina ibivugwa mumwandiko				
6	Ibazwa ryo Kuvuga (speaking) . Urugero: teste yo kuvuga muruhame (Oral presentations), gukora ikiganiro mpaka (Debates/discussions)				

IGICE CYA GATATU

11. USHYIZE KU IJANA (%), WUMVA ARI **KUKIHE KIGERO WAKORA NEZA IBI BIKURIKIRA UKOresha URURIMI RW' ICYONGEREZA**. ANDIKA UMUBARE URI HAGATI YA 0 NA 100.

	Bitangoye kandi nkoresha icyongereza, nashobora gukora ibi bikurikira:	Ikigero
1	Nashobora kuvuga neza mucyongereza umwirondoro wabagize umuryango wange	
2	Nashobora gukora presentation mucyongereza nsobanura ingaruka zo kutagira isuku	
3	Nashobora kuvuga neza mucyongereza uko ibiruhuko bishize byagenze	
4	Nashobora gusobanura neza mucyongereza akamaro k'amazi meza	
5	Nashobora gukoresha icyongereza neza mukiganiro mpaka kubyiza byo kuba mumujyi cyangwa kuba mucyaro	
6	Nashobora kwandikira umuntu uvuga icyongereza musobanurira uko ibiruhuko bishize byagenze	
7	Nashobora kwandika neza mucyongereza inyangirakuru isobanura ingaruka zo kutagira isuku	
8	Nashobora kwandika mucyongereza inkuru mvuga kumunsi wange wambere naje gutangiraho ishuri	
9	Nashobora kwandika neza mucyongereza inshamake y'ibyavugiwe munama y'umuyobozi wishuri	
10	Nashobora gufata note z'ibyavugiwe munama yabaye mucyongereza	
11	Nashobora gusoma no gusobanukirwa umwandiko uvuga kunshingano z'ababyeyi mumuryango	
12	Nashobora gusoma nkanasobanukirwa numwirondoro uri mucyongereza k'umuririmbyi nkunda	
13	Nashobora gusoma nkanasobanukirwa inyangirakuru y'icyongereza ivuga k'uburezi mu Rwanda	

14	Nashobora gusoma nkasobanukirwa inyandiko y'icyongereza kumiterere y'imisozi y'u Rwanda	
15	Nashobora gusoma nkanasobanukirwa umwandiko w'icyongereza uvuga kukamaro k'amazi meza	
16	Nashobora kumva nkanasobanukirwa ikiganiro k'uruhare rw'abakuze muguteza imbere igihugu	
17	Nashobora kumva kandi nkanasobanukirwa ibivugwa mumakuru y'icyongereza kuri radio Rwanda	
18	Nashobora kumva kandi nkanasobanukirwa ibivugwa mukiganiro kubijyanye n'amashuri kuri radio	
19	Nashobora kumva kandi nkanasobanukirwa ikiganiro kubijyanye n'imiterere y'imisozi y'urwanda	
20	Nashobora kumva kandi nkanasobanukirwa ikiganiro kiri mucyongereza kuri radio Rwanda	

Appendix 2A: Student focus group schedule (English version)

Students focus group schedule

.....

1. How easy do you find your assessment tasks? (Ask students to explain)
2. a) What is the usual focus of instruction in English course? (e.g. grammar or reading)
b) To what extent does the content of classroom assessment match what you focus on during instruction?
3. What about your teachers' classroom assessment methods? How do they vary their assessment forms? (For example use of written tests, oral presentations, short answer questions or true/false etc)
4. What role do you play in deciding on the assessment tasks?
5. What type of feedback do you often get from the teacher?
a) Oral comment, b) Written comment, c) grades d) No feedback
6. How clear are you about what you have to do during assessment tasks?
(Inquire whether the teacher gives them assessment criteria/rubrics before assessment)
7. On a scale from 0 (Not certain at all) to 100 (completely certain), how certain are you that you can do the following in good English. (Encourage students to explain why)
a) I can make an oral discussion in good English explaining the importance of literacy (ability to read, write and arithmetic)?
b) I can write a letter in good English explaining the importance of literacy (ability to read, write and arithmetic)?
c) I can read and understand a written text about the importance of literacy (ability to read, write and arithmetic)?
d) I can listen and understand a radio broadcast on the importance of literacy (ability to read, write and arithmetic)?

Appendix 2B: Students focus group schedule (Kinyarwanda version)

Ikiganiro n' abanyeshuri

.....

1. Iyo mubazwa mwishuri mwisomo ry' icyongereza (isuzuma bumenyi-assessment) mubona biba byoroshye cyangwa biba bikomeye? (Saba abanyeshuri gusobanura impamvu)
2. a) Ni ibiki mukunda kwibandaho mwishuri iyo muri kwiga icyongereza? (urugero: grammar, gusoma –reading, cyangwa Kuvuga -speaking)
b) Ibyo mwibandaho mwiga ninabyo mubazwaho mugihe cy'isuzuma bumenyi (assessment)?
c) Muri rusanjye nibihe byibandwaho iyo mubazwa.
3. Nubuhe buryo bukunze gukoreshwa na mwarimu iyo ababaza mwishuri (mugihe cy'isuzuma bumenyi: Urugro: Mwandika ibisubuzo, mukora presentation cy debates)?
4. Nuruhe ruhare mugira muguhitamo ibyo mubazwaho n'uburyo mubazwamo mwishuri?
5. Nyuma y'ibazwa, ni ubuhe buryo mumenya ibyo mwakoze neza nibyo mutakoze neza?
a) Mwarimu arabitubwira mumagambo, b) arabitwandikira adusobanurira,
c) Tubirebera kumanota twabonye , d) abivuga muri rusange d) Ntabwo tujya tubimenya,
6. Nikuruhe rugero muba mwumva neza icyo mugomba gukora mugihe cy' ibazwa?
(Abanyeshuri basobanure niba mwarimu abaha amabwiriza yumvikana y'uko bagomba gukora)
7. Mumibare kuva kuri 0 (ivuga ko uzi ko utabishobora na gato) kugeza kuri 100 (ivuga ko uzineza ko wabishobora neza rwose), hitamo ikigero wumva washobora gukora ibi bikurikira ukoresha icyongereza kiza. -(sabako batanga impamvu bahisemo umubare)
A) Nashobora kuvuga icyongereza neza nsobanura akamaro ko kumenya gusoma/ ko kwiga
b) Nashobora kwandika neza mucyongereza inyandiko nsobanura akamaro ko kumenya gusoma
c) Nashobora gusoma kandi nkumva neza umwandiko wanditse mucyongereza uvuga akamaro ko kumenya gusoma
d) Nashobora kumva kandi nkasobanukirwa neza ikiganiro kiri mucyongereza kuri radio kivuga akamaro ko kumenya gusoma.

Appendix 3A: Students’ assent form (English version)

Student assent form

My name is Viateur Ndayishimiye, a second year PhD research student at The University of York in the United Kingdom. I am writing to invite you to participate in my research entitled: “*Classroom-based assessment and its relationship with the students’ self-efficacy: The case of English language learning in Rwandan lower secondary schools*”. This research is conducted in Kigali city and Southern province schools and your school has been selected to take part in the research.

This study on classroom-based assessment practices of teachers of English in Rwandan lower secondary schools and its effects on the students’ beliefs about their ability to speak and write in English. If you participate in this study, you will be given a questionnaire and asked to answer some questions regarding the assessment practices that you experience in the classroom in the English course. You will also be requested to give your answers on questions regarding your level of confidence for speaking, reading, listening and writing in English. You may also be part of the focus group members who will be requested to meet with the investigator for 15 to 20 minutes interview. In order to maintain confidentiality and protect your identity, no names either of you or your school will appear in any recorded data. Your head teacher, your teacher as well as your parent or your guardian, are all aware of this study. Your participation in this study is not obligatory and you can decide to withdraw at any time.

For any queries, please feel free to contact me at vn553@york.ac.uk or on my Tel. number +447542399011. You can also contact my supervisor Dr Bill Soden at bill.soden@york.ac.uk Or Dr Paul Wakeling who is a member of Education Ethics Committee at education-research-administrator@york.ac.uk. You can also contact the chairperson of the CHMS – IRB (0788490522) or the deputy chairperson (0783340040). If you are willing to participate, please sign your name.

1	confirm that I have read and understood the information given to me about the above named research and I understand that this will involve me taking part as I described above.	✓
2	I understand that I am free to decline participation and can withdraw my data at any point before or during data collection	✓
3	I understand that the information I provide will be used solely for the purpose of the research	✓

Student’s name:.....

Student’s signature:..... Date...../...../.....

Researcher’s signature:..... Date...../...../.....

Appendix 3B: Students' assent form (Kinyarwanda version)

Kwemera kwitabira ubushakashatsi

Nitwa Ndayishimiye Viateur nkaba ndi mumwaka wa kabiri wamasomo y' icyiciro cya gatatu cya kaminuza (PhD) muri iniverisite ya York mu Bwongereza. Iyi nyandiko ni ikumenyesha ubushakashatsi buzakorera mukigo wigaho no kugushishikariza kubwitabira. Ubu bushakashatsi bufite inyito: *“Classroom-based assessment and its relationship with students' self-efficacy: The case of English language learning in Rwandan lower secondary schools”* bukaba buzakorera mumashuri yo muntara y' amajyepfo n'umujyi wa Kigali.

Ubu bushakashatsi bugamije kureba uburyo abarimu bigisha icyongereza babaza abanyeshuri babo no kureba niba hari ingaruka ubwo buryo bugira kukwiyizera kw'abanyeshuri mukuvuga no kwandika icyongereza. Niba wemeye kujya muri ubu bushakashatsi, uruhare rwawe ni urwo gusubiza ibibazo birebana n'uburyo mwarimu wawe akoresha mukubaza mwishuri ndetse nukuntu wumva wiyizeye mukuvuga, kumva, gusoma no kwandika icyongereza. Ushobora kandi kuzaba umwe mubazaba bagize itsinda rizaganira n'umushakashatsi mugihe cy'iminota hagati ya 15 na 20.

Murwego rwo kudasakaza amakuru, ntaho izina ryawe rizagaragara na hamwe kandi ufite uburenganzira bwo kwanga kujya muri ubu bushakashatsi. Umuyobozi w'ikigo cyawe, umwarimu wawe ndetse n'umubyeyi wawe cyangwa uhagarariye umubyeyi wawe bose bamenyeshajwe kandi bemerako ubu bushakashatsi bukorwa.

Uramutse haricyo ushaka kubaza, wanyohereza ubutumwa bwa email kuri vn553@york.ac.uk cyangwa ukampamagara kuri Tel. +447542399011/ (+250) 789934189. Ushobora kandi kubwoherereza umwarimu unyobora ariwe Dr Bill Soden kuri email ye bill.soden@york.ac.uk cyangwa Dr Paul Wakeling uyobora akanama ka iniversite kareba iby'ubuziranenge mubushakashatsi education-research-administrator@york.ac.uk. Ushobora kandi kuvugisha ukuriye akanama gashinzwe iby'ubuziranenge mubushakashatsi kuri kolegi y'ubuvuzi CHMS – IRB (0788490522) cyangwa umwungirije (0783340040). Niba wemeye kugira uruhare murubu bushakashatsi, andika izina ryawe ushyireho umukono (signature)

1	Nasomye kandi nsobanukirwa neza iby'ubu bushakashatsi	✓
2	Nemeye kubwitabira ubu bushakashatsi kandi ndumvako kubwitabira atari itegeko	✓
3	Ndumvako amakuru azatangwa muri ubu bushakashatsi azagirwa ibanga kandi agakoresha gusa murwego rwasobanuwe haruguru	✓

Izina ry'umunyeshuri:.....

Umukono (signature) w'umunyeshuri:..... Itariki:...../...../.....

Umukono w'umushakashatsi..... Itariki:...../...../.....

Appendix 4A. Classroom Observation scheme: Classroom oral interactions

Tch.= To the teacher

WC= To the whole class

St.= To individual student

Gr.St.= To a group of students

Category	Activity	Participant				Total Tallies
		Tch.	WC	St.	Gr.St.	
Oral Questioning	Teachers asks a display question					
	Teacher asks a referential question					
	Teacher asks for more answers					
	Teacher makes a targeted gaze to ask for clarification/ prompt student to answer loudly					
	Teacher explains a grammatical point					
	Teacher explains the meaning of a vocabulary					
	Teacher explains a functional point					
	Teacher explains a point relating to the content of the lesson/assessment					
	Teacher gives instruction/directions on assessment					
	Teacher comments on the student's response to provide further clarification					
	Teacher answers the student question					
	Teacher gives assessment results					
	Teacher praises					
	Teacher criticizes					
	Learners asks a question					
	Learner answers a question					
	Learner talks to another learner					
Period of silence						

Appendix 4B. Classroom Observation scheme: Assessment forms used in the classroom

		Form of Assessment	Tallies	Total Tallies
Performance-based	Speaking	Oral interviews of students		
		Student oral discussion/debate		
		Student oral presentation		
		Oral description		
		Story telling		
		Oral translation		
		Role play		
	Writing	Paragraph writing		
		Essay writing		
		Letter writing		
		Writing summaries		
		Editing		
Paper and Pencil based	Written	Note taking		
		Short-answer items		
		Sentence-completion items		
		True-false items		
		Matching		
		Multiple choice		

	Other practices	Tallies	Total Tallies
Assessment instruction	Rubrics		
	Grading standards		
Student Organization	Student self-assessment		
	Student peer assessment		
	Individual student work		
	Students working in groups		
Feedback	Student peer feedback		
	Collective verbal teacher feedback		
	Individualized verbal teacher feedback		
	Written collective teacher feedback (on blackboard)		
	Individualized written teacher feedback		
	Test results only (marks)		
Source of assessment	Items developed by teacher		
	prepared with other teachers		
	Textbook-embedded items		
	Items from external exam papers		

Appendix 5: Teacher questionnaire

School code: .../

Script No:...

PART I: Personal information

Please cross in the box (X) or write your answer in the provided space.

Q1. Your gender

Male

Female

Q2. Your age group

18-25

26-35

36-50

Above 50

Q3 How long have you been teaching English? Please write the number of Years of English teaching experience

Q4. Which category best describes your highest level of education?

Secondary education

Bachelor's degree

Master's degree

Q5. How often have you had training in assessment techniques in the last five years?

Never

Once

Twice

Three times

More than three times

PART II: Assessment routines

Put X to choose your answer

Q6. At what instructional period is your performance assessment focused (Example of performance assessment: students' oral presentations, debates, essay writing, etc)

	Frequency			
	Usually	Sometimes	Rarely	Never
1. At the end of an instructional unit				
2. At any time during the instruction				
3. In the middle of an instructional unit				
4. At the beginning of an instructional unit				

Q7. At what instructional period is your paper-and-pencil assessment normally focused? (assessment such as written multiple choices, fill-in questions, short answer questions, etc)

	Frequency			
	Usually	Sometimes	Rarely	Never
1. At the end of an instructional unit				
2. At any time during the instruction				
3. In the middle of an instructional unit				
4. At the beginning of an instructional unit				

Q8. How often in a week is the following assessed in your classroom?

	More than three times a week	Two to three times a week	Once a week	Once every two weeks
1. Speaking				
2. Writing				
3. Reading				
4. Listening				
5. Grammar				
6. Vocabulary				
7. Pronunciation				
8. Student class attendance				
9. Student class participation				
10. Student overall effort				

Q9. Please indicate how often the following methods of providing feedback are used in your classroom assessment

	Usually	Sometimes	Rarely	Never
1. Student peer feedback				
2. Collective verbal teacher feedback				
3. Individualized verbal teacher feedback				
4. Written collective teacher feedback (on blackboard)				
5. Individualized written teacher feedback				
6. Test results only (marks)				

PART III: Forms of Assessment used in the classroom

Q 10. How often do you use these assessment forms? Cross (x) to answer

	Usually (e.g. More than three times a week)	Sometimes (e.g. Two to three times a week)	Rarely (e.g. Once a week or once every two weeks)	Never (you think it is essential but you never use it)
1. multiple choice				
2. matching/ true-false				
3. short-answer items				
4. sentence-completion				
5. paragraph writing				
6. essay writing				
7. Summary writing				
8. poem writing				
9. Student Interviews				
10. Discussion/debate				
11. Oral presentation				
12. role play				
13. reading Comprehension				
14. Book reading				
15. listening to audio				

Q11. How often during assessment are your students involved in the following?

	Usually	Sometimes	Rarely	Never
1. Student self-assessment				
2. student peer assessment				
3. Individual student work				
4. Students working in groups				

Q12. Please indicate how often you use assessment items from the stated sources

	Usually	Sometimes	Rarely	Never
1. questions items developed by myself				
2. Prepared together with other teachers				
3. Textbook-embedded items				
4. Items found on the Internet				
5. Items from external exam papers				

PART IV: Purpose of assessment

Q13. In your classroom assessment, how often do you assess your students for the purpose described in the statements below?

	Usually	Sometimes	Rarely	Never
1. Guide students to set their goals and monitor their own learning progress.				
2. Demonstrate to students how to do self-assessment.				
3. Set the criteria for students to assess their own performance in class.				
4. Provide examples of good self-assessment practice				
5. Measure extent of learning at the end of a lesson or subject.				
6. Determine the degree of accomplishment of a desired learning outcome at the end of a lesson.				
7. Assess the quality of student learning in a class at the end of an instruction.				
8. Make final decision about the level of learning that students achieved at the end of a lesson or subject.				
9. Allow students to discover their strengths and learning difficulties in class.				
10. Provide feedback to students in order to improve their learning process.				
11. Make suggestions to students about how they develop better learning strategies.				
12. Explore effective classroom teaching methods and strategies.				
13. Diagnose areas for improvement of instructional activities.				
14. Identify better learning opportunities for students in class.				
15. Continuously collect learning data from students to improve instructional process.				
16. Rank students based on their class performance to inform other school officials.				
17. Provide information to parents about the performance of their children in school.				
18. Examine how one student performs relative to others in my class.				

Appendix 6A: Consent form for teachers (English version)

Dear Sir/Madam,

My name is Viateur Ndayishimiye, a second year PhD research student at The University of York in the United Kingdom. I am writing to invite you to participate in my research entitled: “*Classroom-based assessment and its effects on students’ self-efficacy: The case of English language learning in Rwanda lower secondary school*”. This research is conducted in Kigali city and Southern province schools and your school has been selected to take part in the research.

About the research

This research is conducted to build a knowledge base on teachers’ assessment practices in English language classrooms in Rwanda lower level of secondary education. It also seeks to investigate why teachers choose to assess the way they assess and how students are affected especially in terms of their self-confidence and beliefs in their own abilities. Targeted participants are teachers who currently teach English in second form and/or third form and their students.

Who participates and how?

You are being invited to participate because you are teaching in second form and/or third form of lower level secondary school which are targeted in this study. As for the process of collecting information, a ten to fifteen minute questionnaire will first be available for you to fill and will be followed by videotaping your teaching sessions. Participants in this study will have up to five hours of their teaching sessions video-taped in the period of one week. I will not be present in classroom during videotaping but a video camera will be fitted in your classroom for each of your allotted 5 hours of English lessons per week. I will also request you to find approximately 20 minutes to participate in a face-to-face interview with me after all sessions will have been recorded for a brief discussion on the questionnaire and videotaped sessions. You will be given opportunity to comment on my records of the interview. Overall, besides the classroom video recording, your participation in the research will take approximately 50 minutes.

Anonymity and data protection

As I conduct this research, I remain conscious of my obligation to respect the privacy and confidentiality of the participating respondents’ information. Information that will be collected by use of video camera will be used for the sole purpose of this research and will in no circumstance be shared by other parties other than me, the researcher and, if need be, my supervisor. No names of the participating teacher or other personal details will appear on any of the documents used in this research be it the questionnaire, interview records or my personal

observation notes. All the collected data will be recorded and kept in password protected files only on my personal computer. Any collected information in form of video, audio or written data will be destroyed after the analysis is completed approximately eighteen months after the completion of data collection process.

Right to withdraw and other ethical considerations

There is no obligation to participate in this study and you have all rights to decline any sort of involvement. You can also withdraw from participating in the research any time before or during data collection by contacting me on email or telephone number provided below. By consenting to participate in this research, you are voluntarily giving me permission to use your information in my study specifically in my doctoral thesis and in my other publications. The whole process of collecting information, its analysis and use will all be done in accordance with ethical guidelines provided by The University of York Research Ethics Committee and the Directorate General of Science, Technology and Research in the Ministry of Education in Rwanda. There is no right or wrong way to behave in classroom and there is neither correct nor wrong answer to the questions asked in the questionnaire or in the interview. For that reason, you are kindly called upon to act in your own usual way and provide answers that reflect most what you believe in.

For any queries, please feel free to contact me at vn553@york.ac.uk or on my Tel. number +447542399011. You can also contact my supervisor Dr Bill Soden at bill.soden@york.ac.uk Or Dr Paul Wakeling who is a member of Education Ethics Committee at education-research-administrator@york.ac.uk. You can also contact the chairperson of the CHMS – IRB (0788490522) or the deputy chairperson (0783340040).

Sincerely

1	I confirm that I have read and understood the information given to me about the above named research and I understand that this will involve me taking part as described above.	✓
2	I understand that I am free to decline participation and can withdraw my data at any point before or during data collection	✓
3	I understand that I will be given the opportunity to comment on a written record of the interview	✓
4	I understand that the information I provide will be used solely for the purpose of the research	✓

Name:

Signature: Date:/...../.....

Appendix 6B: Consent form for teachers (Kinyarwanda version)

Bwana/Madame,

Nitwa Ndayishimiye Viateur nkaba ndi mumwaka wa kabiri wamasomo y' icyiciro cya gatatu cya kaminuza (PhD) muri iniversite ya York mu Bwongereza. Iyi nyandiko ni ibatumira kugira uruhare mubushakashatsi bwange bufite inyito: "Classroom-based assessment and its effects on students' self-efficacy: The case of English language learning in Rwanda lower secondary school". Ubu bushakashatsi bukaba buzakorera mumashuri yo muntara y' amajyepfo n'umujyi wa Kigali ishuri ryanyu rikaba riri muyatoranyijwe kujya mubushakashatsi.

Icyo ubu bushakashatsi bugamije

Ubu bushakashatsi bugamije kumenyekanisha uburyo bukoreshwa n'abarimu b'icyongereza mugihe cy' isuzumabumenyi (assessment) ritangirwa mu ishuri mubihe bisanzwe by'amasomo yo mukiciro kibanza cy'amashuri yisumbuye. Bugamije kandi gusuzuma impamvu abarimu b'icyongereza bahitamo bakanakoresha uburyo runaka mugihe babaza abanyeshuri babo (assessment). Hazanarebwa kandi niba ikoreshwa ry'uburyo bumwe cyangwa ubundi rigira ingaruka k'uburyo abanyeshuri biyumva mukuba bashobora kuvuga no kwandika icyongereza. Abatumiwe kujya murubu bushakashatsi ni abarimu bigisha isomo ry' icyongereza mu mwaka wa kabiri n'uwa gatatu w'icyiciro cyambere cy'amashuri yisumbuye hamwe n'abanyeshuri babo.

Icyo abazajya mubushakashatsi bazakora

Iyi nyandiko ibohererejwe kuko mwigisha icyongerezamu mumwaka wa kabiri cyangwa uwagatatu w' icyiciro cyambere cy'amashuri yisumbuye. Bikaba ari murwego rwo kubatumira kwitabira ubu bushakashatsi mutanga amakuru kubijyanye n'imibarize/isuzuma bumenyi (assessment) mu isomo ry'icyongereza. Igikorwa cyo gukusanya amakuru kirimo ibice bitatu. icyambere kirebana n'amakuru yanditse aho musabwa nibura iminota 15 yo kuzuzanya ibibazo byanditse. Hazakurikiraho gufata amashusho n'amajwi mugihe cy'amasaha nibura ane mwigisha icyongereza. Nyuma yo kuzuzanya ibibazo no gufata amashusho, umushakashatsi azabasaba kandi ko muganira mugihe cy'iminota 15 na 20. Nyuma y'ibi bice bitatu muzahabwa umwanya wo kumva ibyafashe mumajwi no kureba ibyafashwe mumashusho kugirango musuzume niba byafashwe uko byatanze. Muri rusange, hazakenerwa iminota isaga 50 yo Kwitabira ubushakashatsi.

Kugira ibanga ibisubizo no kudahishura imyirondoro y'abazajya mubushakashatsi

Ubu bushakashatsi buzakorwa muburyo bwubahirije amategeko yo kurinda no kudasakaza amakuru yatanze n'abitabiriye ubushakashatsi. Ari amakuru azafatwa na kamera ndetse n'ibisubizo byanditse hamwe n'ibiganiro tuzagirana byose bizabikwa neza kandi ntibizabonwa n'undi muntu uretse njyewe nk'umushakashatsi ndetse n'umwarimu wanjye unyobora murubu

bushakashatsi igihe bibaye ngombwa. Ntamyirondoro (nk'amazina) y'umwarimu, y'umuyobozi, y'umunyeshuri cyangwa izina ry'ikigo bizagaragara na hamwe mumakuru azafatwa yaba ayanditse cyangwa ayafashwe na kamera. Amakuru yose azinjizwa mumashini ahantu hafunzwe n'umubare w'ibanga mukwirinda ko hari undi wayabona. Biteganyijwe ko amakuru yose haba ari mubyanditswe cyangwa amajwi bizasenywa nibimara gusuzumwa nyuma y'amezi cumi n'umunani.

Kujya no kuva mubushakashatsi

Nubwo abarimu, abayobozi b'ibigo ndetse n'abanyeshuri bo kubigo byatoranyijwe basabwa gufasha umushakashatsi kugera kumakuru nyayo yafasha kurangiza ubushakashatsi neza, ukwitabira ntabwo ari itegeko. Ashobora kuvamo igihe icyo aricyo cyose mbere cyangwa mugihe cyo gutanga amakuru. Yakoresha email yanjye cyangwa agahamagara kuri numero zanjye zatanzwe hasi amenyesha ko adashaka kwitabira. Kwemera kujya mubushakashatsi bivuzeko uwitabiriye ahaye njyewe umushakashatsi uburenganzira bwo gukoresha amakuru ampaye muryandiko yange isoza icyiciro cya gatatu cya Kaminuza. Ibizakorwa byose bizakorwa hisunzwe amabwiriza agenga ubushakashatsi kuri iniversite ya York ndetse n'ayo mikigo cyo muri minisiteri y'uburezi mu Rwanda gishinzwe iby'ubumenyi, ikoranabuhanga n'ubushakashatsi. Ntagisubizo cyiza cyangwa kibi. Ikingenzi nuko uwitabiriye ubushakashatsi atanga amakuru akurikije uko atekereza.

Haramutse haricyo uwitabiriye ashaka kubaza, yanyohereza ubutumwa bwa email kuri vn553@york.ac.uk cyangwa akampamagara kuri Tel. +447542399011/ (+250) 789934189. Ashobora kandi kubwoherereza umwarimu unyobora ariwe Dr Bill Soden kuri email ye bill.soden@york.ac.uk cyangwa Dr Paul Wakeling uyobora akanama ka iniversite kareba iby'ubuziranenge mubushakashatsi education-research-administrator@york.ac.uk. Ashobora kandi kuvugisha ukuriye akanama gashinwe iby'ubuziranenge mubushakashatsi kuri kolegi y'ubuvuzi CHMS – IRB (0788490522) cyangwa umwungirije (0783340040).

Murakoze

1	Nasomye kandi ndumva neza icyo ubushakashatsi bugamije kandi ko mbigiramo uruhare	✓
2	Ndumva neza ko kwitabira ubushakashatsi Atari itegeko kandiko navamo	✓
3	Ndumva neza ko nzahabwa umwanya wo gusubira kumva no gusuzuma amakuru	✓
4	Ndumva neza ko amakuru ntanze azakoreshwa gusa kumpamvu z'ubu bushakashatsi	✓

Izina:

Umukono: Itariki:/...../.....

Appendix 7A: Consent form for Head teachers (English Version)

Dear Sir/Madam,

My name is Viateur Ndayishimiye, a second year PhD research student at The University of York in the United Kingdom. I am writing to invite you to participate in my research entitled: “*Classroom-based assessment and its effects on students’ self-efficacy: The case of English language learning in Rwanda lower secondary school*”. This research is conducted in Kigali city and Southern province public schools and your school has been selected to take part in the research.

About the research

This research is conducted to build a knowledge base on teachers’ assessment practices in English language classrooms in Rwanda lower level of secondary education. It also seeks to investigate why teachers choose to assess the way they assess and how students are affected especially in terms of their self-confidence and beliefs in their own abilities. Targeted participants are teachers who currently teach English in second form and/or third form and their students.

Who participates and how?

This invitation letter is sent to you because you are the head teacher of a school that has the lower level secondary with 2nd and 3rd forms. Direct participants from your school are 2 teachers who teach in second form and/or third form and their students. As for the process of collecting information, teachers and students will be invited to complete a ten to fifteen minute questionnaire. The student’s questionnaire will be used to elicit information on students’ self-efficacy for speaking and writing English. On the other hand, teachers will be asked to provide information on the purpose and methods that they use in their everyday classroom assessment and how they provide feedback to students. Also, up to five hours of each teacher teaching sessions will be video-taped in the period of one week. To avoid distraction in classroom, I will not be in classroom during the teaching. A video camera will be fitted in the classroom for each allotted 5 hours of English lessons per week. I will also request teachers to find approximately 20 minutes to participate in a face-to-face interview with me after all sessions will have been recorded, for a brief discussion on the questionnaire and videotaped sessions. Participants will be given opportunity to comment on my records of the interview. Overall, besides the classroom video recording, each teacher’s participation in the research will take approximately 50 minutes.

Anonymity and data protection

As I conduct this research, I remain conscious of my obligation to respect the privacy and confidentiality of the information of participating respondents. Information that will be collected by use of video camera will be used for the sole purpose of this research and will in

no circumstance be shown or shared with other parties other than me, the researcher and, if need be, my supervisor. No names of the participating school, head teacher, teacher or student or other personal details will appear on any of the documents used in this research be it the questionnaire, interview records or my personal observation notes. All the collected data will be recorded and kept in password protected files only on my personal computer. Any collected information in form of video, audio or written data will be destroyed after the analysis is completed approximately eighteen months after the completion of data collection process.

Right to withdraw and other ethical considerations

Although participation in this research is highly appreciated and would help answer some of current questions regarding the teaching and learning of English language in Rwanda, there is no obligation to participate and you have all rights to decline any sort of involvement. You can also withdraw from participating in the research any time before or during data collection by contacting me on email or telephone number provided below. By consenting to participate in this research, you are voluntarily giving me permission to use anonymised data from your school in my study specifically in my doctoral thesis and in my other publications. The whole process of collecting information, its analysis and use will all be done in accordance with ethical guidelines provided by The University of York Research Ethics Committee and the Directorate General of Science, Technology and Research in the Ministry of Education in Rwanda. There will be no right or wrong way to behave in classroom and there is neither correct nor wrong answer to the questions asked in the questionnaire or in the interview. Participants will be kindly called upon to act in their own usual way and provide answers that reflect most what they believe in.

For any queries, please feel free to contact me at vn553@york.ac.uk or on my Tel. number +447542399011. You can also contact my supervisor Dr Bill Soden at bill.soden@york.ac.uk or Dr Paul Wakeling who is a member of Education Ethics Committee at education-research-administrator@york.ac.uk. You can also contact the chairperson of the CHMS – IRB (0788490522) or the deputy chairperson (0783340040).

Sincerely,

1	I confirm that I have read and understood the information given to me about the above named research and I understand that this will involve my school taking part as described above.	✓
2	I confirm that the named research is entirely within the norm and my consent gives right to the named researcher to proceed with data collection	✓
3	I understand that my school, teachers and students’ identity will be protected by use of unidentifiable codes and/or pseudonym and that the gathered will be used solely for the purpose of this research	✓
Name of the head teacher: Name of the school:..... Signature: Date:/...../.....		

Appendix 7B: Consent form for Headteachers (Kinyarwanda version)

Bwana/Madame,

Nitwa Ndayishimiye Viateur nkaba ndi mumwaka wa kabiri wamasomo y' icyiciro cya gatatu cya kaminuza (PhD) muri iniversite ya York mu Bwongereza. Iyi nyandiko ni ibatumira kugira uruhare mubushakashatsi bwange bufite inyito: "Classroom-based assessment and its effects on students' self-efficacy: The case of English language learning in Rwanda lower secondary school". Ubu bushakashatsi bukaba buzakorera mumashuri yo muntara y' amajyepfo n'umujyi wa Kigali.

Icyo ubu bushakashatsi bugamije

Ubu bushakashatsi bugamije kumenyekanisha uburyo bukoreshwa n'abarimu b'icyongereza mugihe cy' isuzumabumenyi (assessment) ritangirwa mu ishuri mubihe bisanzwe by' amasomo yo mukiciro kibanza cy'amashuri yisumbuye. Bugamije kandi gusuzuma impamvu abarimu b'icyongereza bahitamo bakanakoresha uburyo runaka mugihe babaza abanyeshuri babo (assessment). Hazanarebwa kandi niba ikoreshwa ry'uburyo bumwe cyangwa ubundi rigira ingaruka k'uburyo abanyeshuri biyumva mukuba bashobora kuvuga no kwandika icyongereza. Abatumiwe kujya murubu bushakashatsi ni abarimu bigisha isomo ry' icyongereza mu mwaka wa kabiri n'uwa gatatu w'icyiciro cyambere cy'amashuri yisumbuye hamwe n'abanyeshuri babo.

Icyo abazajya mubushakashatsi bazakora

Iyi nyandiko ibohererejwe kuko muyoboye ikigo gifite icyiciro cyambere cy'amashuri yisumbuye byumwihariko umwaka wa kabiri n'uwa gatatu. Abazagira uruhare murubu bushakashatsi ni abarimu bigisha icyongereza mumwaka wa kabiri n'uwagatatu kukigo muyoboye hamwe n'abanyeshuri babo. Bazasabwa kuzaza ibibazo byanditse bivuga kubijyanye n'imibarize umwarimu akoresha mu ishuri ndetse n'ikigero abanyeshuri biyumvaho ko bashobora kuvuga no kwandika icyongereza. Abarimu nabo bazasubiza ibibazo bijyanye n'uburyo bakoresha mugihe cy'ibazwa/isuzuma bumenyi mw' ishuri ndetse n'impamvu babukoresha. Hazanafatwa kandi amashusho n'amajwi y'abarimu bari kwigisha mugihe cy'amasaha ane cyangwa atanu yigishwa mu isomo ry'i Cyongereza mucyumweru. Ntabwo umushakashatsi azaba ari mu ishuri mugihe amajwi n'amashusho bizaba bifatwa murwego rwo kwirinda kurangaza. Umushakashatsi kandi azagirana ikiganiro cy'iminota hagati ya 15 na 20 n'abarimu bazaba bafashwe amashusho bigisha. Azaganira kandi n'abanyeshuri mu itsinda rigizwe n'abanyeshuri bageze kuri batandatu mugihe cyingana n' iminota hagati ya 15 na 20. Muri rusange, bizatwara iminota isaga 50 kumwarimu n'umunyeshuri uzaba witabiriye ubushakashatsi asubiza ibibazo byanditse hamwe n'ibiganiro.

Kugira ibanga ibisubizo no kudahishura imyirondoro y'abazajya mubushakashatsi

Ubu bushakashatsi buzakorwa muburyo bwubahirije amategeko yo kurinda no kudasakaza amakuru yatanze n'abitabiriye ubushakashatsi. Ari amakuru azafatwa na kamera ndetse n'ibisubizo byanditse hamwe n'ibiganiro tuzagirana byose bizabikwa neza kandi ntibizabonwa

n'undi muntu uretse n'umushakashatsi ndetse n'umwarimu wanjye unyobora murubu bushakashatsi igihe bibaye ngombwa. Ntamyirondoro (nk'amazina) y'umwarimu, y'umuyobozi, y'umunyeshuri cyangwa izina ry'ikigo bizagaragara na hamwe mumakuru azafatwa yaba ayanditse cyangwa ayafashwe na kamera. Amakuru yose azinjizwa mumashini ahantu hafunzwe n'umubare w'ibanga mukwirinda ko hari undi wayabona. Biteganyijwe ko amakuru yose haba ari mubanditswe cyangwa amajwi bizasenywa nibimara gusuzumwa nyuma y'amezi cumi n'umunani.

Kujya no kuva mubushakashatsi

Nubwo abarimu, abayobozi b'ibigo ndetse n'abanyeshuri bo kubigo byatoranyijwe basabwa gufasha umushakashatsi kugera kumakuru nyayo yafasha kurangiza ubushakashatsi neza, ukwitabira ntabwo ari itegeko. Ashobora kuvamo igihe icyo aricyo cyose mbere cyangwa mugihe cyo gutanga amakuru. Yakoresha email yanjye cyangwa agahamagara kuri numero zanjye zatanzwe hasi amenyesha ko adashaka kwitabira. Kwemera kujya mubushakashatsi bivuzeko uwitabiriye ahaye n'umushakashatsi uburenganzira bwo gukoresha amakuru ampaye munyandiko yange isoza icyiciro cya gatatu cya Kaminuza. Ibizakorwa byose bizakorwa hisunzwe amabwiriza agenga ubushakashatsi kuri iniversite ya York ndetse n'ayo mikigo cyo muri minisiteri y'uburezi mu Rwanda gishinzwe iby'ubumenyi, ikoranabuhanga n'ubushakashatsi. Ntagisubizo cyiza cyangwa kibi. Ikingenzi nuko uwitabiriye ubushakashatsi atanga amakuru akurikije uko atekereza.

Haramutse haricyo uwitabiriye ashaka kubaza, yanyohereza ubutumwa bwa email kuri vn553@york.ac.uk cyangwa akampamagara kuri Tel. +447542399011/ (+250) 789934189. Ashobora kandi kubwoherereza umwarimu unyobora ariwe Dr Bill Soden kuri email ye bill.soden@york.ac.uk cyangwa Dr Paul Wakeling uyobora akanama ka iniversite kareba iby'ubuziranenge mubushakashatsi education-research-administrator@york.ac.uk. Ashobora kandi kuvugisha ukuriye akanama gashinwe iby'ubuziranenge mubushakashatsi kuri kolegi y'ubuvuzi CHMS – IRB (0788490522) cyangwa umwungirije (0783340040).

Murakoze,

1	Nasomye kandi ndumvaneza iby'ubu bushakashatsi	✓
2	Ndemeza ko ubu bushakashatsi bwakorwa mukigo nyoboye	✓
3	Ndumvako amakuru azatangwa n'abarimu n'abanyeshuri bitabira ububushakashatsi azagirwa ibanga kandi agakoreshwa gusa murwego rwasobanuwe haruguru	✓

Izina ry' umuyobozi w' ikigo:

Izina ry' ikigo.....

Umukono: Itariki:/...../.....

Appendix 8A: Consent form for parents (English version)

Dear Sir/Madam,

My name is Viateur Ndayishimiye, a second year PhD research student at The University of York in the United Kingdom. I am writing to invite you to participate in my research entitled: “*Classroom-based assessment and its effects on students’ self-efficacy: The case of English language learning in Rwanda lower secondary school*”. This research is conducted in Kigali city and Southern province public schools and your school has been selected to take part in the research.

Aims of the research and the role of participants

This research is conducted to build a knowledge base on teachers’ assessment practices in English language classrooms in Rwanda lower level of secondary education. It also seeks to investigate why teachers choose to assess the way they assess and how students are affected especially in terms of their self-confidence and beliefs in their own abilities. It is expected that the results of the study will help improve the teaching and learning of English in Rwandan secondary schools. Targeted participants are teachers who currently teach English in second form and/or third form and their students. Participants in this study will be teachers and students of the second and/or third forms of the selected schools. As for the process of collecting information, teachers and students will be invited to complete a ten to fifteen minute questionnaire. The student’s questionnaire will be used to elicit information on students’ self-efficacy for speaking and writing English. Also, up to five hours of each teacher teaching sessions will be video-taped for four to five hours in the period of one week. To avoid distraction in classroom, the investigator will not be in classroom during the teaching. A video camera will be fitted in the classroom for each hour of English lesson. The investigator will also hold a 15 to 20 minutes interview with a group of six to eight students to discuss about the methods of assessment in their classroom and how they feel about writing and speaking English. Participants will be given opportunity to comment on records of the interview.

Anonymity and data protection

Information that will be collected by use of video camera will be used for the sole purpose of this research and will in no circumstance be shown or shared with other parties other than me, the researcher and, if need be, my supervisor. No names of the participating school, head teacher, teacher or student or other personal details will appear on any of the documents used in this research. All the collected data will be recorded and kept in password protected files only on my personal computer. Any collected information in form of video, audio or written data will be destroyed after the analysis is completed approximately eighteen months after the completion of data collection process.

Right to withdraw and other ethical considerations

Although participation in this research is highly appreciated and would help answer some of current questions regarding the teaching and learning of English language in Rwanda, there is no obligation to participate and you have all rights to decline your child’s participation. You can also withdraw from participating in the research any time before or during data collection by contacting me on email or telephone number provided below. By consenting to participate in this research, you are giving me permission to use the collected data in my study specifically in my doctoral thesis and in my other publications. The whole process of collecting information, its analysis and use will all be done in accordance with ethical guidelines provided by The University of York Research Ethics Committee and the Directorate General of Science, Technology and Research in the Ministry of Education in Rwanda.

For any queries, please feel free to contact me at vn553@york.ac.uk or on my Tel. number +447542399011. You can also contact my supervisor Dr Bill Soden at bill.soden@york.ac.uk or Dr Paul Wakeling who is a member of Education Ethics Committee at education-research-administrator@york.ac.uk. You can also contact the chairperson of the CHMS – IRB (0788490522) or the deputy chairperson (0783340040).

Sincerely,

1	I confirm that I have read and understood the information given to me about the above named research and I understand that this will involve my child taking part as described above.	✓
2	I Understand that participation in this study is not obligatory and I consent for participation of my child.	✓
2	I understand that my school, teachers and students’ identity will be protected by use of unidentifiable codes and/or pseudonym and that the gathered will be used solely for the purpose of this research	✓

Name of parent/guardian:

Name of the student:.....

Signature of parent/guardian: Date:

...../...../.....

Researcher’s

signature:.....Date:...../...../.....

Appendix 8B: Consent form for parents (Kinyarwanda version)

Bwana/Madame,

Nitwa Ndayishimiye Viateur nkaba ndi mumwaka wa kabiri wamasomo y' icyiciro cya gatatu cya kaminuza (PhD) muri iniversite ya York mu Bwongereza. Iyi nyandiko ni ibamenyesha ubushakashatsi buzakorwa mukigo umwana wanyu cyangwa uwo muhagarariye yigaho. Ni iyo kubasaba kandi uburenganzira bw' uko umwana wanyu cyangwa uwo muhagarariye yitabira ubu bushakashatsi. Ubu bushakashatsi bufite inyito: "Classroom-based assessment and its effects on students' self-efficacy: The case of English language learning in Rwanda lower secondary schools" bukaba buzakorwa mumashuri yo muntara y' amajyepfo n'umujyi wa Kigali.

Icyo ubu bushakashatsi bugamije n' icyo abazabwitabira bazakora

Ubu bushakashatsi bugamije kureba uburyo bukoreshwa n' abarimu b' icyongereza mugihe cy' isuzumabumenyi (assessment) ritangirwa mu ishuri mubihe bisanzwe by' amasomo yo mukiciro kibanza cy' amashuri yisumbuye. Bugamije kandi gusuzuma impamvu abarimu b' icyongereza bahitamo bakanakoresha uburyo runaka mugihe babaza abanyeshuri babo (assessment). Hazanarebwa kandi niba ikoreshwa ry' uburyo bumwe cyangwa ubundi rigira ingaruka k' uburyo abanyeshuri biyumvamo ubushobozi bwo kuvuga no kwandika icyongereza. Abatumiwe kujya murubu bushakashatsi ni abarimu bigisha isomo ry' icyongereza mu mwaka wa kabiri n' uwa gatatu w' icyiciro cyambere cy' amashuri yisumbuye hamwe n' abanyeshuri babo. Biteganyijwe ko ibizava murubu bushakashatsi bizafasha mukunoza imyigishirize y' icyongereza mumashuri yisumbuye.

Abazagira uruhare murubu bushakashatsi ni abarimu bigisha icyongereza mumwaka wa kabiri n' uwagatatu hamwe n' abanyeshuri babo. Abarimu n' abanyeshuri bazasabwa kuzuba ibibazo byanditse bivuga kubijyanye n' imibarize umwarimu akoresha mu ishuri ndetse n' ikigero abanyeshuri biyumvaho ko bashobora kuvuga no kwandika icyongereza. Hamwe hazanafatwa kandi amashusho n' amajwi y' abarimu bari kwigisha mugihe cy' amasaha ane cyangwa atanu yigishwa mu isomo ry' icyongereza mucyumweru. Ntabwo umushakashatsi azaba ari mu ishuri mugihe amajwi n' amashusho bizaba bifatwa murwego rwo kwirinda kurangaza. Umushakashatsi kandi azagirana ikiganiro cy' iminota hagati ya 15 na 20 n' abanyeshuri bamwe mu itsinda rigizwe n' abanyeshuri bageze kuri batandatu. Muri rusange, bizatwara iminota isaga 50 kumwarimu n' umunyeshuri uzaba witabiriye ubushakashatsi asubiza ibibazo byanditse hamwe n' ibiganiro.

Kugira ibanga ibisubizo no kudahishura imyirondoro y' abazajya mubushakashatsi

Ubu bushakashatsi buzakorwa muburyo bwubahirije amategeko yo kurinda no kudasakaza amakuru yatanze n' abitabiriye ubushakashatsi. Ari amakuru azafatwa na kamera ndetse n' ibisubizo byanditse hamwe n' ibiganiro tuzagirana byose bizabikwa neza kandi ntibizabonwa n' undi muntu uretse njyewe nk' umushakashatsi ndetse n' umwarimu wanjye unyobora murubu

bushakashatsi igihe bibaye ngombwa. Ntamyirondoro (nk'amazina) y'umunyeshuri cyangwa izina ry'ikigo yigaho bizagaragara na hamwe mumakuru azafatwa yaba ayanditse cyangwa ayafashwe na kamera. Amakuru yose azinjizwa mumashini ahantu hafunzwe n'umubare w'ibanga mukwirinda ko hari undi wayabona. Biteganyijwe ko amakuru yose haba ari mubyanditswe cyangwa amajwi bizasenywa nibimara gusuzumwa nyuma y'amezi cumi n'umunani.

Kujya no kuva mubushakashatsi

Nubwo ababyeyi b'abanyeshuri bo kubigo byatoranyijwe gukorerwamo ubushakashatsi basabwa gufasha umushakashatsi kugera kumakuru nyayo yafasha kurangiza ubushakashatsi neza, ukwitabira ntabwo ari itegeko. Uwitabiriye ashobora kuvamo igihe icyo aricyo cyose mbere cyangwa mugihe cyo gutanga amakuru. Yakoresha email yanjye cyangwa agahamagara kuri nimeru zanjye zatanzwe hano amenyesha ko adashaka kwitabira. Gushyira umukono kuriyi nyandiko bivuzeko muhaye nnyewe umushakashatsi uburenganzira bwo gukoresha amakuru nzakusanya munyandiko yange isoza icyiciro cya gatatu cya Kaminuza. Ibizakorwa byose bizakorwa hisunzwe amabwiriza agenga ubushakashatsi kuri iniversite ya York ndetse n'ayo mukigo cyo muri minisiteri y'uburezi mu Rwanda gishinzwe iby'ubumenyi, ikoranabuhanga n'ubushakashatsi.

Haramutse haricyo mushaka kubaza, mwanohereza ubutumwa bwa email kuri vn553@york.ac.uk cyangwa mukampamagara kuri Tel. +447542399011/ (+250) 789934189. Mushobora kandi kubwoherereza umwarimu unyobora ariwe Dr Bill Soden kuri email ye bill.soden@york.ac.uk cyangwa Dr Paul Wakeling uyobora akanama ka iniversite kareba iby'ubuziranenge mubushakashatsi education-research-administrator@york.ac.uk. Mushobora kandi kuvugisha ukuriye akanama gashinzwe iby'ubuziranenge mubushakashatsi kuri kolegi y'ubuvuzi CHMS – IRB (0788490522) cyangwa umwungirije (0783340040).

Murakoze,

1	Nasomye kandi nsobanukirwa iby'ubu bushakashatsi	✓
2	Ndumvako kujya mubushakashatsi atari itegeko kandi ndemezako umwana mugarariye wiga kuriki kigo yitabira ubushakashatsi.	✓
3	Ndumvako amakuru azatangwa muri ubu bushakashatsi azagirwa ibanga kandi agakoreshwa gusa murwego rwasobanuwe haruguru	✓

Izina ry'umubyeyi cyangwa umuhagarariye:

Izina ry'umunyeshuri ahagarariye:.....

Umukono: Itariki:/...../.....

Umukono w'umushakatsi:.....Itariki:/...../.....

Appendix 9: Sample teacher interview transcript (Interview with Teacher JS)

Researcher: May I ask if you assessed your students in this week?

JS: I assess... it's normally after every lesson. I have to assess my students (inaudible)... there is an exercise prepared according to the lesson I have delivered

Researcher: Ok you assess after every lesson

JS: yeah after every lesson I have to assess

Researcher: If you assess after every lesson, what is the aim of your assessment?

JS: The main purpose of every assessment is to make sure that the lesson is well delivered or not and in case the students fail the assessment you have to give once again the lesson and to make sure that they understand well. Before... before the lesson there is a percentage the teacher wants to reach that percentage is shown by the assessment

Researcher: Now about the use of rubrics. How often do you use them?

JS: 'grille de correction' ?

Researcher: yes ... right 'grille de correction'

JS: those are rubrics?

Researcher: yes they are rubrics

T: We don't use them so often ...

Researcher: Why don't you use them?

JS: ...we often ask students to talk about what they think about the teaching and learning of English in general but regarding assessment, we don't use those

Researcher: (inaudible) ...some teachers say that they write them on the questionnaire (inaudible) but those are short instructions explaining how their work will be marked. Saying ABCD criteria have to be met ... that you will get 10/10 if you meet these or those criteria. Do you ever use them?

JS: ehm No. I don't use them

Researcher: I saw in the video that you use oral questioning in the classroom. Are there any other assessment methods that you use for your classroom assessment?

JS: eh other methods are group works. They work in groups to discuss a given exercise, usually written and you go and correct them. It's usually short: 5 to 7 short questions. You go have a quick look on what they do and marking it.

Researcher: Those group assessments, group works, how often in a week do you use them?

JS: On average is 70%. Because they are advantageous in that students do a lot of things in a short time. In a classroom of 40 students, you divide them in 5 groups, you just use a short time to each group. That's what we often use like 70%. They are not time consuming like individual works

Researcher: So those individual and group works are all written. Are there any other methods you use?

JS: : Oral presentations are also often used. You give a question to a group of students, they work on it and you select one student to come in front of the class and present.

Researcher: Between written and oral assessment, what's used more often than the other?

JS: It's written

Researcher: That's what you use often?

T: What is most frequently used is written assessment [...] because oral students say 'we have not prepared' you'll notice that no ideas ... oral ...it's not used often because you have to give them work to bring home, take enough time ...

Researcher: Ok

JS: yes they do written everyday

Researcher: eh so oral presentation requires preparation on the side of the teachers and to...

JS: Both students and teachers

Researcher: Both students and teachers ... now one may wonder... when do you decide that it is time to assess?

JS: I assess immediately after the lesson

Researcher: So it's an established routine that after the lesson you have to assess?

JS: yeah or at the beginning of the lesson you have to assess whether the last lesson is well understood. That's at the beginning, of course in the middle when you are teaching, you also give oral assessment

Researcher: But in general...

JS: In general I assess at the end

Researcher: When your students are working on assessment tasks, what do you do?

JS: within the classroom?

Researcher: within the classroom yes

T: I have to supervise them go through their groups to check whether they are working, because sometimes when you give them work and you don't monitor them, they venture in their own stories in their native language. Some students may also be lazy; some may think that because they are many in the group, some will just sit while others do the work. (inaudible) they may feel it is good time to talk about their own issues in that case one has to be careful because if you are not careful, group work may be a total loss of time.

Researcher: Ok...and how often do you engage your students in group work...I thing we have answered that already... Do you sometimes engage them in peer assessment?

JS: I often take their exercise books and I ask them to correct each other

Researcher: Ok they do that?

JS: Yes

Researcher: Is that often or occasionally and why?

JS: ehm it's occasionally. Because some of them tend to give free marks. You do it when you see that there is not enough time left to go and mark each student.

Researcher: Talking about feedback. Now they've worked on their activities, they submit them to you. What kind of feedback do you provide to them?

JS: The usual feedback is marks. It's marks to tell them how well they worked. When they fail then you know the lesson has not been understood well. You then know that assessment at the end of the lesson is important.

Researcher: I saw in the video that sometimes you ask students to clap for correct answers. Is that done often in your classroom and why?

JS: yeah it's it's done often because it's an encouragement

Researcher: Ok the last question is about the source of your assessment activities. Where do you get them?

JS: The sources are from books especially the written ones but oral depend on how students participate. Of course you prepare questions but others depend on the participation of students.

Researcher: (inaudible) so it's textbook in general...

JS: yeah

Researcher: That's all I had to ask you today thank you very much for your time

JS: Thank you

Appendix 10: Sample transcript from the student focus group (Group B – translated from Kinyarwanda)

Researcher: I thank you all! As I told you earlier this discussion is going to be about assessment, in relation to the way you are assessed in the classroom in the English language subject. Quiz..., tests..., all those kind of activities that the teacher uses to assess you, whether they are marked or not marked. That is what I want us to talk about. Now the first question that I want to ask you is: In general, how do you find your assessment activities? Do you think they are easy or difficult? ... For example when I was a student in secondary school, assessment activities in Math and Physics subjects were always difficult. I hardly found them easy! So in your case...in English subject do you think the assessment activities that you do are easy or difficult? Generally speaking...

S1: In general, assessment tasks are easy. However, sometimes it is difficult to understand what needs to be done. You may be able to read the text but fail to understand it.

Researcher: Ok anyone else with a comment on that?

S2: Actually the biggest challenge is to understand the teacher's method. When the teacher changes the method of assessment which is familiar to us, it becomes difficult to us. We are only familiar with questions starting with "what is..."

Researcher: Ok...it becomes difficult for you when she or he changes the methods?

S2: Yes.

Researcher: Right...(waiting for another student to comment)...others what do you think? Ok...now about the activities, to what extent is the content of the assessment activities related to the learning content?

S (chorus): ...usually it always about what we study in the classroom

Researcher: So you all seem to agree that the content of the assessment activities is always related to the learning content.

S (Chorus): yes/right

Researcher: Right...so my question then is: What is the focus of your learning content?

S1: Verbs

S3: Tenses

Researcher: Ok ...verbs and tenses... anything else?

S4: Nouns

Researcher: That is grammar ... right? How about essay writing? Or writing stories, letters...how often do you do that?

S (chorus): No, we don't do it.

Researcher: You don't do it?

S1: No

Researcher: How about speaking? Like learning how to do debates etc...?

S2: We do it. Every Friday a debate is always scheduled

Researcher: Ok...and that is every week?

S2: It is once a month. The last Friday of every month between students of different class levels

Researcher: I see...how about in your own classroom? Does it happen that the teacher can notify you will learn about speaking? Say ...that you will learn about how to organise an oral presentation in the classroom?

S5: No... that is very rare...we don't do it.

Researcher: Ok...so it's about grammar and assessment is usually focused on grammar as well.

S(chorus): yes

Researcher: Right...in connection to that...may I know the form of assessment that your teacher of English uses to assess you? We kind of...just talked about it already but in general? Because the teacher may seem to prefer using written based activities, or use graded or non graded reading activities, she or she may also have the habit of giving you activities about writing etc... what does your teacher uses very often?

S1: It's often written activities. But also sometimes he asks us to borrow books from the library, read them and summarise them before doing oral presentation in the classroom.

Researcher: ah that's good. So you do summaries too... Do you do it in groups or individually

S(chorus): Individually

Researcher: Ok... That must be time consuming? How often do you do that? Like weekly...?

S5: No it rare

Researcher: ehm ... now let's look at your participation in the assessment process. Your role for example in determining the time of assessment, the way or method of assessment...in short how are you involved in all that?... or you have no consultation with the teacher at all?

S2: When the assessment time comes, he announces the assessment day then we study. When when we see that the designated assessment day approaches and we are not ready, we ask him to change the date so that we can succeed so well on the assessment.

Researcher: And how about the content of the assessment?

S4: He tells us about the content in advance

Researcher: So you don't... bcause sometimes students may work with the teacher to identify areas of the content that they want to be assessed on. Does it happen here?

S5: Yes... they usually tell us about the learning content on which the assessment will be focused.

Researcher: So can you as student ask your teacher to plan assessment about a given content?

S1: No teachers usually choose the chapter from which assessment questions are formulated.

Researcher: So for example you cannot suggest to the teacher that you want to do oral presentation on a given day.

S3: he decides for us and we do what he wants us to do.

Researcher: Ok...he decides for you. Right now you are done with your assessment. For example you have submitted your assignment or have completed your oral presentation, how do you know about your performance on the assessment?

S4: The teacher gives grades

S3: He comes and gives answers in the classroom

S1: After we finish the assessment activities he marks them. When he comes back he brings marks so that we can see where we had a problem, where we did well and where did not

do well then we write the answers. So that we can do well in case he asks the same questions again.

Researcher: You say that he marks ...grade the activities, how does he do it? Does he write some comments? Write the answers on the black board...how is it done?

S4: He writes the answers on the blackboard

S2: Actually he asks the questions orally and asks students to give the answer and to go and write it on the blackboard

Researcher: Ok ...that's how you know if you got the correct answer

S(chorus): yes

Researcher: When he brings your copies do the marks mean anything to you...does it tell you anything about your performance?

S2: Yes.

Researcher: ok now let's talk about how you would rate yourself from zero to one hundred on how certain you are that you can perform communicative tasks in English. For example you want to talk about the importance of knowing how to write and read, to what extent percent that you can do it well speaking in English? Lets' start from here

S1: For me I am certain 56

Researcher: percent

S1: Yes percent

Researcher: please remember the number you choose because we are going to compare numbers for listening, reading and writing. Now were talking about speaking...in English. To what extent per cent that you are certain you can do it?

S2: 40

S3: 40

S4: 49

*S5:*40

Researcher: In writing now. We suppose you are asked to write half a page in English explaining the importance of knowing how to read and write to someone who does not speak Kinyarwanda for instance, how certain percent do you think you can do it well? We start from here this time.

S5: 49

*S4:*58

*S3:*54

*S2:*52

*S1:*65

Researcher: Ok. We now look at reading. We suppose you have a text in front of you about the importance of knowing how to read and write, how certain are you percent that you can read it and understand well the content?

S1: 70

S2: 50

*S3:*30

*S4:*55

S5: 35

Researcher: Right...that's about reading. Now about listening... We suppose they are speaking in English talking about the importance of knowing how to read and write. How confident are you that you can listen to the radio and understand fully what they are talking about.

*S5:*30

*S4:*47

*S3:*29

*S2:*30

*S1:*45

Researcher: Ok...so overall writing and reading are the two top skills in which you are certain that you can do well. Why is that?

S1: Because we have not reached a good level for speaking and listening

Researcher: In speaking and listening?

S(chorus): yes

Researcher: so this is going to be probably the last question. When you chose the number, on what basis did you base your choice? Did you for example think about the marks that you usually get in English or...?

S(chorus): yes. On the marks

Researcher: Ok...anything else? (...waits for answer)

S2: also based on how we feel.

Researcher: your feelings...?

S2: yes

Researcher: Ok...we come to the end of our discussion. I thank you very much for your participation. As I told you at the beginning, none else will have access to this recording other than me and probably my supervisor. Thank you.

Appendix 11: Sample qualitative data analysis: Codes mining process

Category	Code						
	PL	FR	MR	VN	JS	SF	SL
Purpose of assessment	-Measure students gains	-check level of understanding	-Check if they understood -See if I need to change s.th or repeat lesson	-Let students express their ideas	-see if lesson is well delivered -if a preset success percentage is met -know if you have to repeat lesson	-verify whether my objectives are achieved at end of lesson	-check if the lesson well understood -Repeat immediately if something difficult or Focus on it later
Use of rubrics	-in form of instructions -once a term on test /100	-Instructions on questionnaire - shows item weight -Not shared before	-Prepare assessment scheme first -Helps me know how to mark: whether subtract or add marks -Written on a separate sheet -Not shown to students	- For essay writing: I tell them I want you to write good topic sentence, use a period, etc - Given verbally	-don't use them - because there is no form for that	_ Given in form of instructions and written on question paper. -Given to ensure instructions for every questions are clear	-Only for essay writing (instructions) -For others, stu. Have to discover answers on their own: 'Learner-centered'
Assessment methods	-Oral questioning -Write questions on board -Generally written assessment	-Use chalkboard and stud write answers -Clubs & Debates at end of week -St use library to read books	-Marked Quizzes: ask them to read texts -Talk about rubrics for dictation & Composition -Generally written & grammar focused -Speaking like once a month after stu have read books -Oral takes time	-Written assess.: marked essay: checked clarity, sentence, grammar -Dictation, word reading -Generally written -often use oral questioning for spoken -Never mark speaking tasks -Written usually marked	- often written -oral requiring preparation: students read books and present -oral time consuming for teachers and students	-Quiz , tests & 'assessment': talking to individual students on how they are progressing -Both oral (oral questioning) and written used equally	-Usually written works in groups -Home exercises -Oral discussions -Individual book reading, summary writing and oral presentations done in 3 rd term to prepare for nation. Exam

Time of assessment	-At end of lesson -Done very often	-Use term assessment scheme: usually at beginning and end	- before and after lesson -to see what they know, then what they gained	-Marked given when lesson components are covered -May be at end of a chapter -Unmarked given anytime, especially when st seem to be distracted	-After every lesson -a routine to assess at the end of lesson -also sometimes at the beginning	- Quiz anytime -Test at end of course -Also assess before the lesson / form of revision	- Often at the end of lesson. -often unannounced assessment to measure what stud. retained
Monitoring	-Check if anyone is lagging behind -Check if they cheat	-I monitor, see if they are busy doing the work	- go to groups to guide and motivate	-Walk around clarifying instructions	-supervise them -see if they are not distracted -ensure no time is lost: students not focusing on task	-As every teacher, I invigilate: talk to every group to see what's going on, see their difficulties	-I observe: monitor and check if there is anything difficult
Group works	-Pair and group used -Individual works when marked	-'Exactly always because I know the beauty of it'	- rarely give individual works -often use groups -no peer assess. stu not used to it. Gave up on it	-national exam is individual and written-based. So I do the same. No waste of time in groups, or time-consuming speaking -hardly used for student to explain to each other -Sometimes peer-assess given for multiple choices.	-use them at 70% because they are -not time consuming like individual where teacher needs to mark every student -requires good monitoring to avoid st distraction and waste of time. -Peer-assessment rare as st tends to give free marks	-Group discussions often used as 'public speaking' -peer-assessment done for unmarked works: just as encouraged by competence-based curriculum: every student is a teacher to each other'	-used at the end of every lesson -Group discussions -Peer assessment not used: To avoid conflict between students 'given our history of Rwanda ' : Risk of pairing up stud. whose families are not good neighbours

Feedback	-On the blackboard -Students write answers -Individual are rare	-Give them marks -Correct collectively on blackboard -Applause	-give marks first then talk to them -never write on their papers -have bad experience of a teacher accused of harassment -like telling them about behaviour	-Usually collectively orally on the board -Always add comments on paper. Because some students may not follow during oral collective feedback	-marks and collective correction on the board -Also applause for good answer for encouragement	-give positive oral feedback for oral questioning -For written and marked: give marks + comments and also correct on board interacting with students	-mark papers and give marks (to individual or groups) -Difficult questions in groups corrected on blackboard together
Source of tasks	-mainly books, also -Create them & use internet	-Usually books but -Also sometimes internet	- books, syllabus but also my notebooks	-books from the library but –also consult internet sometimes	-Written tasks taken from books -Oral questions depending on how students participate	-Personal own books as school is poor, Past papers (in F 3) -Also internet	-Formulate them as I teach, also –use notebooks from my past studies, also sometimes books

Appendix 12: Instruments and data collection methods used for each research question

Research question (RQ)	Method & data	Instrument
RQ1: What are the classroom assessment practices of teachers of English and how do the students perceive these assessment practices?		
<i>RQ 1.1 What forms of assessment do the teachers use in their classroom?</i>	- 66 teachers & 1258 students survey - 7 teacher interviews & 4 student focus groups - 7 Classroom observations	- Questionnaire - Interview & focus group schedules - Observation schedule
<i>RQ 1.2 For what purpose and at what time do the teachers use assessment forms in the classroom</i>	- 66 teachers & 1258 students survey	- Questionnaire
<i>RQ 1.3 What is the content focus and source of assessment tasks used in the classroom?</i>	- 7 teacher interviews & 4 student focus groups	- Interview & focus group schedules
<i>RQ 1.4 What methods of providing feedback do the teachers use?</i>	- 66 teachers & 1258 students survey - 7 teacher interviews & 4 student focus groups - 7 Classroom observations	- Questionnaire - Interview & focus group schedules - Observation schedule
<i>RQ1.5 What are the students' perceptions of the teachers' classroom assessment practices?</i>	- 1258 students survey	- Questionnaire
RQ2: To what extent is the students' self-efficacy for the four English language skills related to the teachers' assessment practices?		
<i>RQ 2.1 What is the level of the students' self-efficacy for the four English language skills?</i>	- 1258 students survey	- Questionnaire
<i>RQ 2.2 How is the students' self-efficacy for the four English language skills related to the students' perceptions of the teachers' assessment practices?</i>	- 66 teachers & 1258 students survey - 7 teacher interviews & 4 student focus groups	- Questionnaire - Interview & focus group schedules
<i>RQ 2.3 How is the students' self-efficacy for the four English language skills related to the teachers' assessment practices?</i>	- 7 Classroom observations	- Observation schedule

Appendix 13: Results from Spearman test of correlation between the teachers' use of performance assessment forms related to reading and the students' self-efficacy for reading in English.

13.A Results from the analysis of data from teachers

			Students' self-efficacy in reading	Use of performance assessment related to reading
Spearman's rho	Students' self-efficacy in reading	Correlation Coefficient	1.000	-.057
		Sig. (2-tailed)	.	.694
		N	51	51
	Use of performance assessment related to reading	Correlation Coefficient	-.057	1.000
		Sig. (2-tailed)	.694	.
		N	51	51

13.B Results from the analysis of data from students

			Students' self- efficacy in reading	Use of performance assessment related to reading
Spearman's rho	Students' self-efficacy in reading	Correlation Coefficient	1.000	.072*
		Sig. (2-tailed)	.	.011
		N	1244	1244
	Use of performance assessment related to reading	Correlation Coefficient	.072*	1.000
		Sig. (2-tailed)	.011	.
		N	1244	1244

*. Correlation is significant at the 0.05 level (2-tailed).

Appendix 14: Results from Spearman test of correlation between the teachers' use of performance assessment forms related to listening and the students' self-efficacy for listening in English.

14.A Results from the analysis of data from teachers

			Use of performance assessment related to listening	Students' self-efficacy in listening
Spearman's rho	Use of performance assessment related to listening	Correlation Coefficient	1.000	.316
		Sig. (2-tailed)	.	.068
		N	34	34
	Students' self-efficacy in listening	Correlation Coefficient	.316	1.000
		Sig. (2-tailed)	.068	.
		N	34	34

14.B Results from the analysis of data from students

			Students' self-efficacy in listening	Use of performance assessment related to listening
Spearman's rho	Students' self-efficacy in listening	Correlation Coefficient	1.000	.022
		Sig. (2-tailed)	.	.433
		N	1232	1232
	Use of performance assessment related to listening	Correlation Coefficient	.022	1.000
		Sig. (2-tailed)	.433	.
		N	1232	1232

References

- Adamson, B. (2004). Fashions in language teaching methodology. In A. Davies & E. Catherine (Eds.), *The handbook of applied linguistics* (pp. 623-642). Oxford: Blackwell Publishing.
- Ahn, H. (2015). Assessing proficiency in the National English Ability Test (NEAT) in South Korea: A critique of a government's approach to testing English proficiency. *English Today*, 31(1), 32-42.
- Airasian, P.W. (2005). *Classroom assessment. Concept and application* (5th ed.). New York, NY: McGraw-Hill.
- Akyel, A. (2000). Collaboration to explore teaching: A case study. *TESL Canada Journal*, 18(1), 58-74.
- Allal, L. & Lopez, M. L. (2014). Teachers' professional judgement in the context of collaborative assessment practice. In C. Wyatt-Smith, V. Klenowski & P. Colbert (Eds.), *Designing Assessment for Quality Learning* (pp. 151-166). New York, NY: Springer.
- Alderson, J. C. & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14 (2), 115-129.
- Alderson, J.C & Banerjee, J. (2001). Language testing and assessment: State-of-the-art review. *Language Teaching*. 34, 213-236.
- Alkharusi, H. (2008). Effects of classroom assessment practices on students' achievement goals. *Educational Assessment*. 13(4), 243-266.
- Alkharusi, H., Alnafri, S. & Alnabhani, H. (2014). Classroom assessment: teacher practices, student perceptions, and academic self-efficacy beliefs. *Social Behavior and Personality*, 42(5), 835-856.
- Alkharusi, H. (2013). Canonical correlational models of students' perceptions of assessment tasks, motivational orientations, and learning strategies. *International Journal of Instruction*, 6(1), 21-38.

- Altschuler, S.J. & Schmautz, T. (2006). No Hispanic student left behind: The consequences of “high stakes” testing. *Children & Schools*, 28(1), 5–14, <https://doi.org/10.1093/cs/28.1.5>
- Alsop, A. & Ryan, S. (1996). *Making the most of fieldwork education: A practical approach*. Berlin: Springer Science + Business Media.
- Alvarez, L., Ananda, S., Walqui, A., Sato, E., & Rabinowitz, S. (2014). *Focusing formative assessment on the needs of English language learners*. San Francisco: WestEd.
- Ananda, S. & Rabinowitz, S.N. (2002). *The high stakes of high-stakes testing*. Retrieved on 28 March 2015 from <https://www.wested.org/resources/the-high-stakes-of-high-stakes-testing>.
- Anderson, L.W. (2003). *Classroom assessment: Enhancing the quality of teacher decision making*. Mahway: Lawrence Erlbaum Associates Inc.
- Andrade, H. (2000). Using rubrics to promote tinkering and learning. *Educational Leadership*, 57(5), 13-18. Retrieved on 16 August 2017 from <http://www.ascd.org/publications/educational-leadership/feb00/vol57/num05/toc.aspx>
- Andrade, H.; Wang, X.; Du, Y.; & Akawi, R. L.(2009). Rubric-referenced self-assessment and self-efficacy for writing. *Educational Theory and Practice Faculty Scholarship*. Paper 13. Retrieved on 20 February 2016 from http://scholarsarchive.library.albany.edu/etap_fac_scholar/13
- Andrade, H. & Du, Y. (2005). Student perspectives on rubric-referenced assessment. *Practical Assessment and Research Evaluation*, 10(3). retrieved on 10 August 2017 from <http://pareonline.net/pdf/v10n3.pdf>
- Andrews, R. (2003). *Research questions*. London: Continuum.
- Anjamshoa, L. & Sadighi, F. (2015). The importance of motivation in second language acquisition. *International Journal on Studies in English Language and Literature (IJSELL)*, 3(2), 126-137.

- Arnold, J. (2007). Self-concept and the affective domain in language learning. In F. Rubio (Ed.), *Self-esteem and foreign language learning*, (pp 13-29). Cambridge: Cambridge Scholars Publishing.
- Arter, J. & Mctighe, J. (2001). *Scoring rubrics in the classroom: Using performance criteria for assessing and improving student performance*. LA, CA: Corwin Press Inc.
- Artino, A. R. (2012). Academic self-efficacy: From educational theory to instructional practice. *Perspect Med Educ*, 1, 76–85, DOI 10.1007/s40037-012-0012-5.
- Azadi, G. & Gholami, R. (2013). Feedback on washback of EFL tests on ELT in L2 classroom. *Theory and Practice in Language Studies*, 3(8), 1335-1341.
- Baadte, C. & Schnotz, W. (2013). Feedback effects on performance, motivation and mood: Are they moderated by the learner's self-concept? *Scandinavian Journal of Educational Research*, doi: 10.1080/00313831.2013.781059
- Babbie, E. (2013). *The basics of social research* (6th ed.). Belmont: Wardsworth Cengage Learning.
- Bachman, L. (1990). *Fundamental considerations in language testing*. New York: Oxford University Press.
- Bailey, K. (1999). *Washback in language testing*. Princeton, NJ: ETS.
- Bailey, K. (2008). *Methods of social research*. New York: The Free Press.
- Baker, P. & Ellice, S. (2011). *Key terms in discourse analysis*. London: Continuum International Publishing Group.
- Bandura, A. (1978). The self system in reciprocal determinism. *American Psychologist*, 33(4), 344-358.
- Bandura, A. (1984). Recycling misconceptions of perceived self-efficacy. *Cognitive Therapy and Research*. 8(3), 231-255.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

- Bandura, A. (1995). *Self-Efficacy in changing societies*. Cambridge: Cambridge University Press.
- Bandura, A. (1997). *Self-Efficacy: The exercise of control*. New York, NY: Worth Publishers.
- Bandura, A. (1999). Exercise of personal and collective efficacy in changing societies. In A. Bandura (Ed.), *Self-efficacy in changing societies* (pp. 1–45). New York, NY: Cambridge University Press.
- Bandura, A. (2006). Guide for Constructing Self-Efficacy Scales. In F. Pajares & T. Urdan (Eds.), *Self-efficacy beliefs of adolescents*, (pp. 307-337). Greenwich: Information Age Publishing.
- Banks, S. (2012). *Classroom Assessment: Issues and Practices* (2nd ed.). Long Grove, Waveland Press.
- Barker, G., & Graham, S. (1987). Developmental study of praise and blame as attributional cues. *Journal of Educational Psychology*, 79, 62-66.
- Barnes, D. (1990). Language in the secondary classroom. In D. Barnes, J. Britton & M. Torbe (Eds.), *Language, the learner and the school* (4th ed.), (pp 9-88). Portsmouth: Cook Publishers.
- Barton, A.C., Atweh, B., Borba, M.C., Gough, N., Keitel, C., Vistro-Yu, C. & Vithal, R. (2008). *Internationalisation and globalisation in mathematics and science education*. Dordrecht: Springer Science and Business Media.
- Bashejja, D. (2014). Use of written English by students and staff educated in French: A case of institutions of higher learning in Kigali. *Rwanda Journal of Education*, 2(2), 9-16
- Beleghezadeh, S. & Masoun, A. (2013). The effect of self-assessment on EFL Learners' Self-Efficacy. *TESL Canada Journal /Revue TESL Du Canada*, 31(1), 42-58.
- Bernhardt, V. (2013). *Data analysis for continuous school improvement* (2nd ed.). New York, NY: Routledge.

- Berryman, S.E. (2000). *Hidden challenges to education systems in transition economies*. Washington, DC: The World Bank.
- Black, P. J.; Harrison, C.; Lee, C.; Marshall, B. & Wiliam, D (2004). Working inside the black box: assessment for learning in the classroom. *Phi Delta Kappan*, 86 (1), 8-21.
- Black, P. & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan*, 80(2), 139–48.
- Black, P. & Wiliam, D. (1998b). Assessment and classroom learning. *Assessment in Education: Principles, Policy & Practice*, 5(1), 7-74.
- Black, P. & Wiliam, D. (2009). Developing the theory of formative assessment, *Educational Assessment and Evaluation*, 21 (1), 5-31.
- Bless, C., Higson-Smith, C. & Kagee, (2007). *Fundamentals of social research methods: An African Perspective*. Kenwyn: Juta & Company Ltd.
- Blindley, G. (2001). Outcomes-based assessment in practice: some examples and emerging insights. *Language Testing*, 18 (4) 393–407.
- Borg, W. R. (1963). *Educational research: An introduction*. London: Longman.
- Borg, W. R. & Gall, M. D. (1979). *Educational Research: An Introduction* (3rd ed.). London: Longman.
- Brace, I. (2013). *Questionnaire design: How to plan, structure and write survey material for effective market research*. London: Kogan Page Publishers.
- British Council Rwanda (2015). *Supporting teachers' English through mentoring (STEM)*. Retrieved on 26 January 2018 from <https://www.britishcouncil.rw/programmes/education/supporting-teachers-english-through-mentoring-stem>
- Brookhart, S. M. (1994). Teachers' grading: Practice and theory. *Applied Measurement in Education*, 7(4), 279-301.
- Brookhart, S. M. (1998). Determinants of student effort on schoolwork and school-based achievement. *The Journal of Educational Research*, 91(4), 201–208.

- Brookhart, S.M. (2013). *How to create and use rubrics for formative assessment and grading*. Alexandria: ASD Publications.
- Brookhart, S. M. (2010). *Formative assessment strategies for every classroom*. Alexandria: ASCD.
- Brookhart, S.M & Durkin, D.T. (2003). Classroom assessment, student motivation, and achievement in high school social studies classes, *Applied Measurement in Education*, 16(1), 27-54, doi: 10.1207/S15324818AME1601_2
- Brown, H. D. (2004). *Language assessment: Principles and classroom practices*. New York : Longman.
- Brown, H. D. & Abeywickrama, P. (2010). *Language assessment: principles and classroom practices*. Harlow: Pearson Education.
- Brown, J. & Hudson, T. (1998). *The alternatives in language assessment*. *TESOL Quarterly* 32(4). <https://doi-org.libproxy.york.ac.uk/10.2307/3587999>
- Bryman, A. (2012). *Social Research Methods* (4th ed.). Oxford, Oxford University Press.
- Butler, S. M. & McMunn, N. D. (2006). *Teacher's guide to classroom assessment: Understanding and using assessment to improve student learning*. San Francisco: John Wiley & Sons, Inc.
- Butler, Y.G. (2015). Parental factors in children's motivation for learning English: a case in China, *Research Papers in Education*, 30(2), 164-191, doi: 10.1080/02671522.2014.891643
- Cargan, L. (2007). *Doing social research*. Plymouth: Rowman & Littlefield Publishers, Inc.
- Central Intelligence Agency - CIA (n.d). *The world factbook, Africa: Rwanda*. retrieved on 02 January, 2018 from <https://www.cia.gov/library/publications/the-world-factbook/geos/rw.html>
- Chapelle, A.C. (1999). Validity in language assessment. *Annual Review of Applied Linguistics*. 19, 254-272. <https://doi-org.libproxy.york.ac.uk/10.1017/S0267190599190135>

- Chapelle, C. A. & Douglas, D. (2006). *Assessing language through computer technology*, Cambridge: Cambridge University Press.
- Chaudron, C. (1988). *Second language classrooms: Research on teaching and learning*. Cambridge: CUP.
- Chen, H. (2007). *The relationship between efl learners' self-efficacy beliefs and English performance*. Retrieved on 21 January 2016 from <http://diginole.lib.fsu.edu/islandora/object/fsu%3A254251>
- Chen, L. & Sun, Y. (2015) Teachers grading decision makin: Multiple influencing factors and Methods. *Language Assessment Quarterly*, 12 (2), 213-233
- Cheng, H. & Dornyei, Z. (2007). The use of motivational strategies in language instruction: The case of efl teaching in Taiwan. *Innovation in Language Learning and Teaching*, 1(1), 153-174.
- Cheng, L & Watanabe, J. Y. (Eds.), (2004). *Washback in language testing. Research contexts and methods*. London: Lawrence Earlbaum Publishers.
- Cheng, L. & Wang, X. (2007). Grading, Feedback, and Reporting in ESL/EFL Classrooms. *Language Assessment Quarterly*, 4(1), 85–107.
- Cheng, L. (1996). *Washback effect of public examination change on classroom teaching*. Hong Kong: University of Hong Kong.
- Cheng, L., Rogers, T. & Hsu, H. (2004). ESL/EFL instructors' classroom practices: Purposes, methods, and procedures. *Language Testing*, 21 (3), 360–389.
- Cheng, L., Rogers, W.T & Wang, X. (2008). Assessment purposes and procedures in ESL/EFL classrooms. *Assessment and Evaluation in Higher Education*, 33(1), 9-32, <https://doi.org/10.1080/02602930601122555>
- Cheng, L., Wu, Y. & Bettney, E. (June 2014) Assessment and motivation: Perspectives from teacher candidates. In S. Van Nuland (Eds.), *58th Yearbook of Teacher Education*. Paper presented at International Council on Education for Teaching 58TH World Assembly. University of Ontario Institute of Technology, Oshawa. June 16-19. (pp 358-371), ICET.

- Chumun, S. (2002). *Classroom assessment in Mauritanian primary schools* (Doctoral dissertation). Retrieved on 19 April 2017 from bura.brunel.ac.uk/bitstream/2438/5154/1/FulltextThesis.pdf
- Cizek, G.J. (2000). Pockets of resistance in the assessment revolution. *Educational Measurement: Issues and Practice*, 19(2), 16-23.
- Clack, R. & Adler, S.E. (2014). *An invitation to social research: How it's done*. Stamford: Cengage Learning.
- Clapham, C. (2000). Assessment and testing. *Annual Review of Applied Linguistics* 20, 147–161. <https://doi-org.libproxy.york.ac.uk/10.1017/S0267190500200093>
- Clotfelter, C.T., Ladd, H. F. & Vigdor, J. L. (2006). Teacher-student matching and the assessment of teacher effectiveness. *Journal of Human Resources*, 41(4), 778-820.
- Cohen, L; Manion, L; & Morrison, K. (2007). *Research methods in education* (6th ed.). London: Routledge.
- Cohen, L; Manion, L; & Morrison, K. (2011). *Research methods in education* (7th ed.). London: Routledge.
- Colton, D. & Covert, R.W. (2007). *Designing and constructing instruments for social research and evaluation*. San Francisco, Jossey-Bass.
- Coniam, D. & Falvey, P. (2007). High-stakes testing and assessment: English language teachers benchmarking. In J. Cummins & C. Davidson (Eds.), *International handbook of English language teaching* (pp. 457- 472). New York: Springer.
- Cooper, B.S. & Gargan, A. (2009). Rubrics in education: Old term, new meanings. *The Phi Delta Kappan*, 91(1), 54-55.
- Coopersmith, S. (1967). *The antecedents of self-esteem*. San Francisco: W. H. Freeman & Co.
- Coronado-Aliegro, J. (2006) *The effect of self-assessment on the self-efficacy of students studying Spanish as a foreign language* (Doctoral dissertation). Pittsburgh: University of Pittsburgh. Retrieved on 14 February from <http://d-scholarship.pitt.edu/9914>

- Cousin, G. (2009). *Researching learning in higher education: an introduction to contemporary methods and approaches*. London: Routledge.
- Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches* (2nd ed.). London: Sage Publications.
- Creswell, J. W. (2005). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research* (2nd ed.). Upper Saddle River, NJ: Pearson.
- Crooks, T. (1988). The impact of classroom evaluation practices on students. *Review of Educational Research*, 58(4), 438-481.
- Cumming, A. (2001). ESL/EFL instructors' practices for writing assessment: specific purposes or general purposes? *Language Testing*, 18 (2), 207-224.
- Curtis, W., Murphy, M. & Shields, S. (2013). *Research and education*. London: Routledge.
- D'Ailly, H. (2003). Children's autonomy and perceived control in learning: A model of motivation and achievement in Taiwan. *Journal of Educational Psychology*, 95(1), 84-96.
- Davison, C. (2004). The contradictory culture of teacher-based assessment: ESL teacher assessment practices in Australian and Hong Kong secondary schools. *Language Testing*, 21 (3), 305-334.
- DeLuca, C., Valiquette, A., Coombs, A., LaPointe-McEwan, D., & Luhanga, U. (2016). Teachers' approaches to classroom assessment: A large-scale Survey, *Assessment in Education: Principles, Policy & Practice*, DOI:10.1080/0969594X.2016.1244514
- Denscombe, M. (2014). *The good research guide: For small-scale social research projects* (5th ed.). Berkshire: Open University Press.
- Denscombe, M. (2008). Communities of practice: A research paradigm for the mixed methods approach. *Journal of Mixed Methods Research* 2 (3), 270-283.
- Diener, E. & Crandall, R. (1978). *Ethics in social and behavioral research*. Chicago: The University of Chicago.

- Dinther, M., Dochy, F. Segers, M. & Braeken, J. (2014). Student perceptions of assessment and student self-efficacy in competence-based education, *Educational Studies*, 40(3), 330-351, DOI: 10.1080/03055698.2014.898577
- Doghonadze, N. (2017). Assessment of writing skills. In N. Doghonadze (eds), *Teaching reading and writing skills in georgia*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Dorman, j., Fisher, D & Waldrup, B. (2006). *Classroom environment, students' perceptions of assessment, academic efficacy and attitude to science: A Lisrel analysis*. Retrieved on 14 January 2017 from <http://core.ac.uk/download/pdf/11039825.pdf>
- Dornyei, Z. (1994). Motivation and motivating in the foreign language classroom. *The Modern Language Journal*, 78(3), 273-284.
- Dornyei, Z. (2001). *Motivational strategies in the language classroom*. Cambridge: Cambridge University Press.
- Dornyei, Z. (2007). *Research methods in applied linguistics*. Oxford: Oxford University Press.
- Drever, E. (2003). *Using semi-structured interviews in small-scale research: A teacher's guide*. (2nd ed.). Glasgow: Scottish Council for Research in Education.
- Dunn, K., Strafford, E. & Marston, C. (2003). *Classroom assessment practices in English and mathematics at Years 5, 7, and 9*. Wellington: New Zealand Council for Educational Research. Retrieved on 12 January 2016 from <http://core.ac.uk/download/pdf/11039825.pdf>
- Dusabemungu, V. (2014, November 14). Rwanda clears French language on its new 2000 Frw & 5000 Rwf Bank notes. *Igihe*. Retrieved on 30 May 2017 from <http://en.igihe.com/news/rwanda-clears-french-language-on-its-new-2000-frw.html>
- Dweck, C. S. (1999). *Self-theories: Their role in motivation, personality and development*. Philadelphia: Psychology Press.
- Dweck, S. (2007). *Mindset: the new psychology of success*. New York: Random House

- Earl, L. (2003). *Assessment as learning: Using classroom assessment to maximise student learning*. Thousand Oaks, CA: Corwin Press.
- Earl, L. (2010). Assessment - A powerful lever for learning. *Brock Education*, 16(1), 1-15.
- Education Policy and Data Centre – EPDC (2014). *National Education Profile-Rwanda*. retrieved on 14 February 2017 from www.epdc.org/sites/default/files/documents/EPDC%20NEP_Rwanda.pdf
- Elliot, C. N., Kratochwill, T. R., Cook, J. L., & Tavers, J. F. (2000). *Educational psychology: Effective teaching, effective learning*. New York: McGraw Hill.
- Ellis, R. (2003). *Task-based language learning and teaching*. Oxford: Oxford University Press.
- Erikson, Erik H. (1994). *Identity and the life cycle*. London : Norton.
- Evans, M. (2013). Analysing qualitative data. In E. Wilson (Ed.), *School-based research: A guide for education students*, (pp. 157-170). London: Sage Publications Ltd.
- Field, A. (2009). *Discovering statistics using SPSS* (3rd ed.). London, Sage Publications Ltd.
- Fisher, R.J. (1993). Social desirability bias and the validity of indirect questioning. *Journal of Consumer Research*, 20 (2), 303–315, <https://doi.org/10.1086/209351>
- Flammer, A. (1995). Developmental analysis of control beliefs. In A. Bandura (Ed.), *Self-efficacy in changing societies*. (pp. 202–231). Cambridge: Cambridge University Press.
- Frey, B.B. & Schmitt, V.L. (2010). Teachers’ classroom assessment practices. *Middle Grades Research Journal*, 5(3), 107–117.
- Fulcher, G. (2012). Assessment literacy for the language classroom. *Language Assessment Quarterly*, 9(2), 113-132. doi: 10.1080/15434303.2011.642041
- Gardner, R. C. (2010). *Motivation and second language acquisition: The Socio-educational Model*. Oxford: Peter Lang.
- Genç, G., Kuluşaklı, E., & Aydın, S. (2016). Exploring EFL learners’ perceived self-efficacy and beliefs on English language learning. *Australian Journal of Teacher Education*, 41(2), 52-68. <http://dx.doi.org/10.14221/ajte.2016v41n2.4>

- Gideon, L. (2012). *Handbook of survey methodology for the social*. New York: Springer Science & Business Media.
- Ghonsooly, B. & M, Elahi (2011). Learners' Self-efficacy in Reading and its relation to Foreign Language Reading Anxiety and Reading Achievement. *Journal of English Language Teaching and Learning Year 53*(217).
- Goodwin, J. (2009). *Research in psychology: Methods and design*. New Jersey: John Wiley & Sons, Inc.
- Gonzales, R. & Aliponga, J. (2011). Classroom assessment preferences of Japanese language teachers in the Philippines and English language teachers in Japan. *MEXTESOL Journal*, 36(1), 1-19.
- Graham, S., & Weiner, B. (1996). Theories and principles of motivation. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 63–84). New York: Simon & Schuster Macmillan.
- Gravetter, F.J. & Forzano, L.B. (2015). *Research methods for the behavioral sciences*. (5th ed.). Stamford, Cengage Learning.
- Green, A. (2014). *Exploring language assessment and testing: Language in action*. Abingdon: Routledge.
- Grinnell, R. M. Jr. & Unrau, Y. A. (2011). *Social foundations of evidence-based practice* (9th ed.). Oxford: Oxford University Press.
- Hall, R. (2008). *Applied social research: Planning, designing and conducting real-world research*. Basingtoke: Palgrave Macmillan.
- Hammond, M. & Wellington, J.J. (2013). *Research methods: The key concepts*. Oxon: Routledge.
- Harlem, W. & Crick, R.D. (2002). A Systematic review of the impact of summative assessment and tests on students' motivation for learning. EPPI – Centre. Retrieved on 13 February 2016 from <http://www.storre.stir.ac.uk/bitstream/1893/19607/1/SysRevImpSummativeAssessment2002.pdf>

- Harris, P. (2005). *Reading in the primary school years*. Sydney: Cengage Learning.
- Hartas, D. (2010). Quantitative research as a method of research in education. in D. Hartas (Ed.), *Educational research and inquiry: Qualitative and quantitative approaches*, (pp 65-81). London: Continuum International Publishing Group.
- Hatch, J.A. (1995). *Qualitative research in early childhood settings*. Westport: Greenwood Publishing Group.
- Hattie, J. & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Hill, K. & McNamara, T. (2012). Developing a comprehensive, empirically based research framework for classroom-based assessment. *Language Testing*, 29(3), 295-420.
- Hill, K. (2017). Understanding classroom-based assessment practices: a precondition for teacher assessment literacy. *Papers in Language Testing and Assessment*, 6(1), 1-17.
- Hillier, Y. (2005). *Reflective teaching in further and adult education* (2nd ed.). London, Continuum.
- Hitchcock, G. & Hughes, D. (2002). *Research and the teacher: A qualitative introduction to school-based research* (2nd ed.). London: Routledge.
- Hsieh, P.H. (2008). Why are college foreign language students' self-efficacy, attitude, and motivation so different? *International Education*, 38(1), 76–94.
- Hsu, H. (2009). *The impact of implementing English proficiency tests as a graduation requirement at Taiwanese universities of technology*. (Unpublished dissertation), University of York, York.
- Inbar-Lourie, O. & Donitsa-Schmidt, S. (2009). Exploring classroom assessment practices: the case of teachers of English as a foreign language. *Assessment in Education: Principles, Policy & Practice*, 16(2), 185–204.
- Jenks, C.J. (2004). The Effects of age, sex, and language proficiency on the self-efficacy of English language learners. *Annual Review of Education, Communication and Language Sciences*, 1, 1-6.

- JeuneAfrique (2016, Août 19). *Afrique centrale : le Rwanda réintègre officiellement la CEEAC.* Retrieved on 23 April 2017 from <http://www.jeuneafrique.com/350335/economie/afrique-centrale-rwanda-reintegre-officiellement-ceac>
- JHA (2014). *Social research methods*. New Delhi: McGraw Hill.
- Johnnie, D. (2012). *Sampling essentials: Practical guidelines for making sampling choices*. Thousand Oaks: Sage Publications, Inc.
- Johnson, R.B. & Onwuegbuzie, A.J (2004). Mixed methods research: A research paradigm whose time has come. *Educational Researcher*, 33(7), 14-26.
- Jonsson, A. (2014). Rubrics as a way of providing transparency in assessment. *Assessment and Evaluation in Higher Education*, 39(7), <https://doi.org/10.1080/02602938.2013.875117>
- Kamanzi, V. (2012, March 08). Guharanira kwesa imihigo bitera abarimu gukopeza abanyeshuri mu bizamini. *Igihe*. Retrieved on 12 January 2015 from <http://www.igihe.com/ikoranabuhanga/ubumenyi/guharanira-kwesa-imihigo-bitera-abarimu-gukopeza-abanyeshuri-mu-bizamini>
- Kanamugire, E. (2017, March 30). REB yaciye undi muvuno nyuma yo gusezerera ba ‘Mentors’ mu Cyongereza. *Igihe*, Retrieved on 20 January 2018 from <http://igihe.com/amakuru/u-rwanda/article/reb-yaciye-undi-muvuno-nyuma-yo-gusezerera-ba-mentors-mu-cyongereza>
- Kaneza, F. (2012). Where English is the new must-have language, but 98% of the people can't speak it. Retrieved on 01 July 2017 from <https://www.worldcrunch.com/culture-society/where-english-is-the-new-must-have-language-but-98-of-the-people-can039t-speak-it>
- Kang’ethe, S.M. (2014) Peer assessment as a tool of raising students’ morale and motivation: The perceptions of the university of Fort Hare Social Work Students. *International Journal of Educational Sciences*, 6(3), 407-413
- Khalk, W.M., Sheir, A.A. & Nabawy, E.E. (2014). Oral questioning technique for developing critical thinking skills in efl classroom. *Educational Sciences Journal*, 22(1)

- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity: International Journal of Methodology*, 47(4), 2025-2047. <http://dx.doi.org/10.1007/s11135-011-9640-9>
- Krysik, J.L & Finn, J. (2013). *Research for effective social work practice*. London: Routledge.
- Kuciel, E.P (2013). Self-efficacy in L2: A research proposal. In K. Piatkowska and K. Okonska (Eds). *correspondences and contrasts in foreign language pedagogy and translation studies*.(pp 31-42), London: Springer International Publishing.
- Kwizera, C. (2011, May 16). Over 6,000 to be admitted in public varsities. *The New Times*. Retrieved on 12 June 2015 from <http://www.newtimes.co.rw/section/read/31196>
- Lado, R. (1961). *Language testing: The construction and use of foreign language tests*. London: Longman.
- Lantolf, J.P. & Poehner, M.E. (2010). Dynamic assessment in the classroom: Vygotskian praxis for second language development. *Language Teaching Research*. 15(1), 11–33.
- Lee, I. (2009). Ten mismatches between teachers' beliefs and written feedback practice. *ELT Journal*, 63(1), 13-22. DOI:10.1093/elt/ccn010
- Leung, C & Mohan B. (2004) Teacher formative assessment and talk in classroom context: Assessment as discourse and assessment of discourse. *Language Testing*, 21 (3), 335-359.
- Leung, C. & Rea-Dickins, P. (2007). Teacher assessment as policy instrument: Contradictions and capacities. *Language Assessment Quarterly*, 4(1), 6–36.
- Lindsay, G. (2010) Ethical considerations and legal issues in educational research. In D. Hartas (Ed.), *Educational research and inquiry: Qualitative and quantitative approaches*. (pp.110-125). London: Continuum International Publishing Group.
- Littlewood, W. & Yu, B. (2011) First language and target language in the foreign language classroom. *Language Teaching*, 44 (1), 64–77.

- Liu, L. (2008). Co-teaching between native and non-native English Teachers: An exploration of co-teaching models and strategies in the Chinese primary school context. *Reflections on English Language Teaching*, 7(2), 103–118. Retrieved on 10 August 2017 from: <http://www.nus.edu.sg/celc/publications/RELT72/103to118liu.pdf>
- Lorin, W. A. (2003). *Classroom assessment: Enhancing the quality of teacher decision making*. London: Routledge.
- Ma, X. (2008). The skills of teacher's questioning in English classes. *International Education Studies*, 1(4), 92-100.
- Mabry, L. (1999). Writing to the rubric: Lingering effects of traditional standardized testing on direct writing assessment. *The Phi Delta Kappa*, 80(9), 673-679.
- Mann, S. (2005). The language teacher's development. *Language Teaching*, 38 (3), 103–118. doi: 10.1017/S0261444805002867
- Marlow, C. (2010) *Research methods for generalist social work*. Belmont: Brooks/Cole Cengage Learning.
- Marsden, P. V. & Wright, D. J. (2010). *Handbook of survey research* (2nd ed.), Bingley, EGP Ltd.
- Marzano, R. (2006). *Classroom assessment & grading that work*. Oroville: Association for Supervision & Curriculum Development-ASCD.
- Martella, R.C.; Nelson, R.J.; Morgan, R.L.; Marchand-Martella, E.N. (2013). *Understanding and interpreting educational research*. New York: Guilford Press.
- Maxwell, A.J. (2005). *Qualitative research design: An interactive approach* (2nd ed.) London: Sage Publications.
- McGreal, C. (2009, January 16). Why Rwanda said adieu to french. *The Guardian*. Retrieved on 25 February 2017 from <https://www.theguardian.com/education/2009/jan/16/rwanda-english-genocide>
- McMillan, J. H. (1997). *Classroom assessment: Principles and practice for effective instruction*. Boston: Allyn & Bacon.

- McMillan, J. H., & Workman, D. J. (1998). *Classroom assessment and grading practices: A review of the literature*. Retrieved on 16 May 2017 from <https://eric.ed.gov/?id=ED453263>
- McMillan, J.H., Myran, S. & Workman, D. (2002). Elementary teachers' classroom assessment and grading practices. *The Journal of Educational Research*, 95(4) 203-213, DOI:10.1080/00220670209596593
- McNamara, T. (2001). Language Assessment as social practice: Challenges for research. *Language Testing*. 18 (4) 333–349.
- McTighe, J. & Ferrara, S. (1998). Assessing learning in the classroom. Student assessment series. *National Education Association*, Washington, DC. Retrieved on 23 July 2916 from <https://files.eric.ed.gov/fulltext/ED429989.pdf>
- Mekonnen, G.T. (2014). EFL classroom assessment: Teachers' practice and teaching techniques adjustment in Ethiopia. *Educational Research and Reviews*, 9(20), 1071- 1089.
- Mehregan, M. (2014). The role of teacher feedback in enhancing learner self-efficacy and motivation in computer-assisted environments. *MEXTESOL Journal*, 38(3), 1-16
- Mertler, C.A. (1998). Classroom assessment practices of Ohio teachers. Conference paper: EDRS. Retrieved on 06 August 2017 from <http://files.eric.ed.gov/fulltext/ED428085.pdf>
- Mills, N. A. (2004). Self-efficacy of college intermediate French students: Relation to motivation, achievement, and proficiency. (Doctoral dissertation, University Microfilms International, US). *Dissertation Abstracts International Section A: Humanities and Social Sciences*, 65 (2-A), 440.
- Mills, N. (2009). Self-efficacy in the standards of foreign language learning instrument. Retrieved on 10 February 2015 from http://works.bepress.com/nicole_mills/28/
- Mills, N., Pajares, F. & Herron, C. (2007). Self-efficacy of college intermediate french students: Relation to achievement and motivation. *Language Learning* 57(3), 417–442.

- Mills, N. (2013). *Self-Efficacy and foreign language education*. Brigham, Brigham Young University
- Milnes, T. & Cheng, L. (2008). Teachers' Assessment of ESL Students in Mainstream Classes: Challenges, strategies, and decision-making. *TESL Canada Journal/Revue TESL du Canada*, 25(2), 49-65.
- MINECOFIN – Ministry of Finance and Economic Planning- (2005). A synthesis of the analyses of the 2002 census of Rwanda. Kigali, Minecofin.
- MINEDUC (2003). Law N° 29/2003 of 30/08/2003 Establishing the organisation and the functioning of nursery, primary and secondary schools. Retrieved on 30 January 2015 from <http://planipolis.iiep.unesco.org/en/2003/law-no-292003-30082003-establishing-organisation-and-functioning-nursery-primary-and-secondary>
- MINEDUC (2010). *Education sector strategic plan*. Kigali, Mineduc.
- MINEDUC (2011). *Curriculum of English language for ordinary level of secondary education in Rwanda*. Kigali: Mineduc.
- MINEDUC (2015a). *National education for all 2015 review*, Kigali: Mineduc.
- MINEDUC (2015b). *2015 education statistical yearbook*. Kigali: Mineduc.
- MINEDUC (2016). *2016 education statistical yearbook*. Kigali: Mineduc.
- Mislevy, R.J., Steinberg, L.S. & Almond, R.G (2002). Design and analysis in task-based language assessment. *Language Testing*, 19(4) 477-696.
- Mitchell, M.L. & Jolley, J.M. (2010). *Research design explained* (7th ed.). New York, NY: Wadsworth, Cengage Learning.
- Moriarty, B. (2014). Research design and the predictive power of measures of self-efficacy. *Issues in Educational Research*, 24(1), 55-66.

- Mucunguzi, R. M. (2018, March 5). Mu mashuri yose ya kaminuza bagiye kujya bakora ikizamini cy'icyongereza. *Igihe*. Retrieved on 13 March 2018 from <http://igihe.com/amakuru/u-rwanda/article/mu-mashuri-yose-ya-kaminuza-bagiye-kujya-bakora-ikizamini-cy-icyongereza>
- Mugisha, I. R. (2017, February 9). Rwanda to make Kiswahili compulsory in schools. *The EastAfrican*. Retrieved on 15 January 2018 from <http://www.theeastafrican.co.ke/news/Rwanda-to-make-Kiswahili-compulsory-in-schools/2558-3807044-xiep8r/index.html>
- Mulkeen, A. (2006). *Teachers for rural schools: A challenge for Africa*. Retrieved on 10 May 2017 from http://www.adeanet.org/adea/biennial2006/doc/document/B2_2_mulkeen_en.pdf
- Multon, K. D., Brown, S. D., & Lent, R. W. (1991). Relation of self-efficacy beliefs to academic outcomes: A meta-analytic investigation. *Journal of Counseling Psychology*, 38, 30-38.
- Munoz, P. & Alvarez, M.E. (2010). Washback of an oral assessment system in the EFL classroom. *Language Testing*, 27(1) 33–49.
- Murray, R.T. (1998). *Conducting educational research: A comparative view*. Westport: Greenwood Publishing Group.
- Mussawy, S.A.J (2009). *Assessment practices: Student's and teachers' perceptions of classroom assessment*. Retrieved on 18 March 2015 from http://scholarworks.umass.edu/cie_capstones/9
- Naderi, S. (2014). The effect of explicit and recast feedback on the intermediate EFL learners' listening self-efficacy. *International Journal of Learning, Teaching and Educational Research*. 2 (1), 30-43.
- Naseri, M & Zaferanieh, E. (2012). The relationship between reading self-efficacy beliefs, reading strategy use and reading comprehension level of Iranian EFL learners. *World Journal of Education*, 2(2), 64-74.

- Natriello, G. (1987). The impact of evaluation processes on students. *Educational Psychologist*, 22 (2), 155-175.
- Ndahiro, A., Gatere, J. & Nkusi, A. (Eds.), (2015). *Rwanda: Rebuilding a nation*. Kampala: Fountain Publishers.
- Newfields, T. (2006). Teacher development and assessment literacy. *Authentic Communication: Proceedings of the 5th Annual JALT Pan-SIG Conference*, 48 - 73.
- Newlyn, D. (2013). Providing exemplars in the learning environment: The case for and against. *Universal Journal of Educational Research* 1(1): 26-32, DOI:10.13189/ujer.2013.010104
- Nkosana, L. (2008). Attitudinal obstacles to curriculum and assessment reform. *Language Teaching Research*. 12 (2), 287–312.
- Nunan, D. & Bailey, K.M. (2009). *Exploring second language classroom research: A comprehensive guide*. Boston: Heinle, Cengage Learning.
- Ntawitonda, J. C. (2014, April 14). *Mu gihe gito icyizere cyo kubaho mu Rwanda cyiyongereyeho imyaka 13. Igihe*. Retrieved on 27 January 2018 from <http://igihe.com/amakuru/u-rwanda/article/mu-gihe-gito-icyizere-cyo-kubaho>
- Odeke, S. (2015, July 25). We need to improve English proficiency across Rwanda. *The New Times*. Retrieved on 10 October 2015 from <http://www.newtimes.co.rw/section/article/2015-07-25/190942>
- Onwuegbuzie, A. J., Bailey, P., & Daley, C. E. (2001). Cognitive, affective, personality and demographic predictors of Foreign Language Achievement. *The Journal of Educational Research*, 94, 3–15.
- Oroujlou, N. & Vahedi, M. (2011). Motivation, attitudes and language learning. *Social and Behavioral Sciences*, 29 (2011) 994 – 1000.
- Ortiz, M. A (2015). The qualitative interview. In F.K. Stage & K. Manning (Eds.), *Research in college context: Research and methods*.(pp. 47-61). London: Routledge.

- Osae, C. (2015, July 22). Francophone to Anglophone: Is Rwanda on the right track? *The New Times*. Retrieved on 19 May 2017 from <http://www.newtimes.co.rw/section/read/190826>
- Outeiral, A. M. C. (2014). Assessment of young language learners: Using rubrics to bridge the gap between praxis and curriculum. *Bellaterra Journal of Teaching & Learning Language & Literature*, 7(1), 52-77.
- Overton, T. (2012). *Assessing learners with special needs: An applied approach* (7th ed.), Brownsville: Pearson.
- Oz, H. (2014). Turkish teachers' practices of *assessment for learning* in the English as a foreign language classroom. *Journal of Language Teaching and Research*, 5(4), 775-785.
- Ozcan, S. (2010). *The effects of asking referential questions on the participation and oral production of lower level language learners in reading classes*. Retrieved on 07 August 2017 from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.634.1528&rep=rep1&type=pdf>
- Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research*, 66 (4), 543-578.
- Pajares, F. (2003). Self-efficacy beliefs, motivation, and achievement in writing: A review of the literature. *Reading & Writing Quarterly*. 19, 139-158,
- Pajares, F. (2002). Gender and perceived self-efficacy in self-regulated learning. *Theory into Practice*, 41(2), 116-125.
- Pajares, F. & Schunk, D.H. (2002). *The development of academic self-efficacy*. Retrieved from <http://www.uky.edu/~eushe2/Pajares/SchunkPajares2001.PDF>
- Pallant, J. (2010). *SPSS Survival manual: A step by step guide to data analysis using SPSS* (4th ed.), Maidenhead : McGraw-Hill, Open UP.

- Panadero, E. & Romero, M. (2014). To rubric or not to rubric? The effects of self-assessment on self-regulation, performance and self-efficacy, *Assessment in Education: Principles, Policy & Practice*, 21(2), 133-148, DOI: 10.1080/0969594X.2013.877872
- Pathak, R.P. (2008). *Methodology of educational research*. New Delhi: Atlantic Publishers & Dist.
- Pearson, P. (2014). Policy without a plan: English as a medium of Instruction in Rwanda. *Current Issues in Language Planning*, 15(1), 39-56.
- Pettifor, J.L. & Saklofske, D. H (2012). Fair and ethical student assessment practices. In C.F Webber, and J.L. Lupton (Eds.), *Leading student assessment*. (pp. 87-106). London: Springer Science + Business Media.
- Phakiti, A. (2006). Theoretical and pedagogical issues in ESL/EFL teaching of strategic reading. *University of Sydney Papers in TESOL*, 1, 19-50.
- Phye, G. D. (1996). *Handbook of classroom assessment: Learning, adjustment, achievement*. London: Academic Press Ltd.
- Pokaew, K. (2004). The roles of attitude and motivation in second and foreign language learning. *Language Institute Journal*, 69(2), 69-76.
- Pollard, A., Triggs, P., Broadfoot, P., Mcness, E. & Osborn, M. (2000). *What pupils say: Changing policy and practice in primary education*. London: Continuum.
- Popham, W. J (2008). *Transformative assessment*. Alexandria: Association for Supervision & Curriculum Development.
- Popham, W. J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice*, 48(1), 4-11.
- Pui, K. L. (2010). *The role of teacher feedback in enhancing student motivation in english school-based assessment*. Hongkong: University of Hongkong.
- Puchta, H. (1999). Beyond materials, techniques and linguistic analysis: The role of motivation, beliefs and identity. *Plenary session at the 33rd International IATEFL Annual conference*, Edinburgh, 64-72.

- Qassim, J. A. (2008). *Teachers' perceptions of current assessment practices in public secondary schools in the state of Qatar*. Hull: University of Hull.
- Randal, S.; Coast, E., Compaore, N. & Antoine, P. (2013). The power of the interviewer. *Demographic Research*, 28 (27), 763-792.
- Read, J. & Hayes, B. (2003). The impact of IELTS on preparation for academic study in New Zealand. In R. Tulloh (Ed.), *International English Language Testing System (IELTS) Research Reports*, 4, (pp. 153-191). Canberra: IDP: IELTS Australia
- Read, J. (2008). Identifying academic language needs through diagnostic assessment. *Journal of English for Academic Purposes*, 7(3), 180-190.
- Rea-Dickins, P. & Germaine, K. (1992). *Evaluation*. Oxford: Oxford University Press.
- Rea-Dickins, P. (2001). Mirror, mirror on the wall: Identifying processes of classroom assessment. *Language Testing*, 18 (4), 429-462.
- Rea-Dickins, P. (2006). Currents and eddies in the discourse assessment: A learning-focused interpretation. *International Journal of Applied Linguistics*. 16(2), 163-188.
- Rea-Dickins, P. (2007). Classroom-based assessment: Possibilities and Pitfalls. In J. Cummins & C. Davison (Eds.), *International handbook of English language teaching*, (pp. 505-520). New York: Springer.
- Rea-Dickins, P. (2008). Classroom-based language assessment. In E. Shohamy & N.H. Hornberger (Eds.), *Encyclopedia of language and education*. Vol. 7, *Language Testing and Assessment*, (pp.257-271). New York: Springer.
- Rea-Dickins, P. & Gardner, S. (2000). Snares and silver bullets: Disentangling the construct of formative assessment. *Language Testing*, 17(2), 215- 243.
- REB (2014, January). Improving quality education. *Rwanda Education Board Newsletter*, 3.
- Reddick, C. (2015). *Alternatives to current medium of instruction policy in Rwandan primary schools*. Harvard Graduate School of Education, Retrieved on 02 January 2018, from <https://wiki.harvard.edu/confluence/display/GlobalEd/Panels>

- Rhodes, R.C., Rosenbusch, M.H. & Thompson, N. (1996). Foreign languages: Instruments, techniques, and standards. In G.D. Phye (eds) *Handbook of classroom assessment: Adjustment, learning and achievement*, (pp 381-415). London: Academic Press Ltd.
- Ritchie, L. (2016) *Fostering self-efficacy in higher education students*. London: Palgrave MacMillan
- Robins, R. W. & Pals J. L. (2002). Implicit self-theories in the academic domain: Implications for goal orientation, attributions, affect, and self-esteem change. *Self and Identity*, 1(4), 313-336.
- Robson, C. (2011). *Real world research* (3rd ed.). Trenton, NJ: John Wiley & Sons.
- Rogier, D. (2014). Assessment literacy: Building a base for better teaching and learning. *English Teaching Forum*, 3rd November. Retrieved on 06 February 2018 from https://americanenglish.state.gov/files/ae/resource_files/etf_52_3_02-13.pdf
- Ross, A.J., Rolheiser, C. & Hogaboam-Gray, A. (2002). Influences on student cognitions about evaluation. *Assessment in Education: Principles, Policy & Practice*, 9(1), 81-91.
- Rotberg, I. C. (2006). Assessment around the world. *Educational Leadership*. 64(3), 58-63
- Rubin, A. & Babbie, E. (2009). *Essential research methods for social work*. (2nd ed.). Belmont: Wardsworth ,Cengage Learning.
- Rust, C. (2002). The impact of assessment on student learning. *Active Learning in Higher Education*, 3(2), 145-158 Retrieved on 20 June 2016 from <http://alh.sagepub.com/cgi/content/abstract/3/2/145>
- Rwirahira, R. (2017, October 22). Struggles stalk English as a second language in Rwanda. *The EastAfrican*. Retrieved on 26 January 2018 from <http://www.theeastafrican.co.ke/rwanda/News/Struggles-English-language-in-Rwanda/1433218-4148974-ephqg4z/index.html>
- Ryumon, A. (2007). Washback of high-stakes assessment, year 12 Japanese. In H. Marriott, T. Moore & R. Spence-Brown (Eds.), *Learning discourses and the discourses of learning*, (pp. 171-185), Clayton: Monash University ePress.

- Saldana, J. (2009). *The coding manual for qualitative researchers*, London: Sage Publications Ltd.
- Samuelson, B.L. & Freeman, S. W. (2010). Language policy, multilingual education, and power in Rwanda. *Language Policy*, 9 (3), 191–215, DOI 10.1007/s10993-010-9170-7
- Sara, K. & Ruilei, L. (2013). Washback of a high-stakes English test in China: Student and teacher perceptions. *Concordia Working Papers in Applied Linguistics-COPAL*, 4, 22-29.
- Sarantakos, S. (2013). *Social research* (4th ed.). Basingtoke: Palgrave Macmillan.
- Saris, W. E. & Gallhofer, I. N (2014). *Design, evaluation, and analysis of questionnaires for survey research*. New Jersey: John Wiley & Sons, Inc.
- Sawafi, O. (2014). *Investigating English teachers' beliefs and practices in relation to the continuous assessment reform in the sultanate of Oman* (Doctoral dissertation). Retrieved on 23 May 2017 from <http://etheses.whiterose.ac.uk/8101>
- Schofield, J.W. (2000). Increasing the generalizability of qualitative research, in R. Gomm, M.Hammersley and P. Foster (Eds.), *Case study method: Key issues, key texts* (pp. 200-225). London: SAGE Publications Ltd.
- Schunk, D. H. (1995). Self-efficacy, motivation, and performance. *Journal of Applied Sport Psychology*, 7(2), 112-137.
- Schunk, D.H (1982). Effects of effort attributional feedback on children's perceived self-efficacy and achievement. *Journal of Educational Psychology*, 74, 548-556.
- Schunk, D. H., & Rice, J. M. (1991). Learning goals and progress feedback during reading comprehension instruction. *Journal of Reading Behavior*, 23(3), 351-364.
- Schunk, D. H., & Zimmerman, B. J. (1997). Social origins of self-regulatory competence. *Educational Psychologist*, 32, 195-208.
- Schunk, D. H. (2003). Self-efficacy for reading and writing: Influence of modeling, goal setting, and self-evaluation. *Reading and Writing Quarterly*, 19, 159-172.

- Schunk, D. H., & Zimmerman, B. J. (2007). Influencing children's self-efficacy and self-regulation of reading and writing through modeling. *Reading and Writing Quarterly*, 23, 7-25.
- Seong, Y. (2011). The teacher's role in classroom-based language assessment. *Columbia University Working Papers in TESOL & Applied Linguistics*, 11(1), 50-52.
- Shams, N. and Tavakori, M. (2014). The effect of peer, self, and traditional assessment on Iranian efl learners' L2 reading comprehension. *Journal of Applied Linguistics and Language Research*, 1(1), 29-44.
- Shapland, N. (2011). *Assessment matters*. Retrieved on 20 December 2016 from <https://app.griffith.edu.au/assessment-matters/docs/design-assessment/principles/transparent>
- Sharma, R.K. (1997). *Sociological methods and techniques*. Delhi: Nice Printing Press.
- Shim, K.N. (2009). *An investigation into teachers' perceptions of classroom-based assessment of English as a foreign language in Korean primary education* (Doctoral dissertation), Retrieved on 12 January 2016 from <https://ore.exeter.ac.uk/repository/handle/10036/67553>
- Shohamy, E. (1993). *The power of tests: the impact of language tests on teaching and learning*. Washington, DC: The National Foreign Language Center at the Johns Hopkins University. Retrieved on 26 May 2015 from <http://files.eric.ed.gov/fulltext/ED362040.pdf>
- Shohamy, E., Donitsa-Schmidt, S. & Ferman, I. (1996). Test impact revisited: Washback effect over time. *Language Testing*. 13, 3, 298-317.
- Shohamy, E. (1997). Testing methods, testing consequences: Are they ethical? Are they fair? *Language Testing*. 14, 3, 340-49.
- Shohamy, E. (2005). The power of tests over teacher: the power of teachers over tests. in D.J. Tedick (Ed.), *Second language teacher education: international perspectives* (pp.101-111). London : Lawrence Erlbaum Associates.

- Sibomana, E. (2014). The acquisition of English as a second language in Rwanda: Challenges and promises. *Rwanda Journal of Education*, 2(2), 19-30.
- Simpson, J. & Muvunyi, E. (2012). Teacher training in Rwanda and the shift to English-medium education. In A. Robertson & R. Jones-parry (Eds.), *Commonwealth Education Partnerships 2012/13*; (pp. 154-157). London: Commonwealth Secretariat.
- Skinner, E., & Greene, T. (2008). Perceived control: Engagement, coping, and development. In T. L. Good (Ed.), *21st Century Education: A Reference Handbook* (pp 121–130). Newbury Park: Sage Publications.
- Smith, J. K., Smith, L. F., Gilmore, A., & Jameson, M. (2012). students' perceptions of reading ability, enjoyment of reading and reading achievement. *Learning and Individual Differences*, 22 (2), 202-206. doi:10.1016/j.lindif.2011.04.010
- Smith, C.D., Worsfold, K., Davies, L., Fisher, R. & McPhail, R. (2013). Assessment literacy and student learning: the case for explicitly developing students 'assessment literacy', *Assessment & Evaluation in Higher Education*, 38(1), 44-60
DOI:10.1080/02602938.2011.598636
- Spolsky, B. (1995). *Measured words*. Oxford: Oxford University Press.
- Statistics Solutions (2013). *Data analysis plan: Spearman rank correlation* [www document]. Retrieved on 1st June, 2017from <http://www.statisticssolutions.com/academicsolutions/memberresources/memberprofile/dataanalysisplantemplates/dataanalysisplanspearmanrankcorrelation>
- Stefani, L. A. J. (1998). Assessment in partnership with learners. *Assessment and Evaluation in Higher Education*, 23(4), 339–350.
- Stefanou, C., & Parkes, J. (2003). Effects of classroom assessment on student motivation in fifth-grade science. *The Journal of Educational Research*, 96(3).
<https://doi.org/10.1080/00220670309598803>
- Stepp-Greany, J. (2004). Collaborative teaching in an intensive Spanish course: A professional development experience for teaching assistants. *Foreign Language Annals*, 137(3), 417-416.

- Stiggins, R.J. (2002). *Assessment crisis: The absence of assessment FOR Learning*. Phi Delta Kappa International. Retrieved on 02 February 2017 from <http://journals.sagepub.com/doi/abs/10.1177/003172170208301010?journalCode=pdka>
- Stiggins, R.J. (1991). Assessment literacy. *The Phi Delta Kappan*, 72 (7), 534-539.
- Supovitz, J. (2009). *Is High-stakes testing working?* Retrieved on 15 July 2015 from <http://www.gse.upenn.edu/review/feature/supovitz>
- Suskie, L. A. (2009). *Assessing student learning: A common sense guide*. San Francisco, CA : Jossey-Bass.
- Tabaro, J. D (2012, August 3). Why Rwandans find it hard to learn English. *The East African*. Retrieved on 21 December 2016 from <http://www.theeastafrican.co.ke/Rwanda/Lifestyle/Why-Rwanda-finds-it-hard-to-learn-English/-/1433242/1470280/-/item/1/-/eyixd6/-/index.html>
- Tante, A. C. (2010). The purpose of English language teacher assessment in the English-speaking primary schools in Cameroon. *LTED*, 13, 27-39. Retried on 10 June 2017 from <http://www.elted.net/volume-13.html>
- Taylor, F. (2013). *Self and identity in adolescent foreign language learning*. Bristol: Multilingual Matters Ltd.
- Teasdale, A. & Leung, C. (2000). Teacher assessment and psychometric theory: A case of paradigm crossing? *Language Testing*. 17 (2), 163–184.
- Thomas, K.M; Usher, E. & Mamaril, N. (2012). *Investigating the relationship between teacher feedback and student self-efficacy*. Retrieved on 20 July 2017 from <https://sites.education.uky.edu/motivation/files/2013/08/ThomasUsherMamaril.pdf>
- Thornbury, S. (1996). *Teachers research teacher talk*. Retried on 08 August 2017 from https://www.researchgate.net/publication/30984586_Teachers_research_teacher_talk
- Tittle, C. K. (1994). Toward an educational psychology of assessment for teaching and learning: theories, contexts, and validation arguments, *Educational Psychologist*, 29, 149–162.

- Topping, K. (1998). Peer assessment between students in colleges and universities. *Review of Educational Research*. 68(3), 249-276.
- Torrance, H. & Pryor, J. (2001). Developing formative assessment in the classroom: Using action research to explore and modify theory. *British Educational Research Journal*, 27(5), 615-631.
- Tsagari, D. & Pavlou, P. (2009). Practices, policies and tensions in classroom-based assessment. *The 6th Annual EALTA Conference*, 4 - 7 June 2009.
- UNESCO (2014). *EFA global monitoring report 2013/4 – teaching and learning: achieving quality for all*. Paris: UNESCO, retrieved on 12 January 2016 from unesdoc.unesco.org/images/0022/002266/226662e.pdf
- University of York (n.d.). *Policy on research degrees*. Retrieved on 06 December 2015 from <https://www.york.ac.uk/research/graduate-school/support/policies-documents/research-degree-policy/#integrity>
- USAID Rwanda (2010). *Assessment report and proposal for an education strategy*. USAID
- Uwishyaka, L. (2015, August 25). Abarimu n’abayobozi b’ibigo babujijwe kuvuga ikinyarwanda igihe cyose bari ku ishuri. *Makuruki*. Retrieved on 30 March 2017 from <http://www.makuruki.rw/IMIBEREHOMYIZA/UBUREZI/article/AbarimunabayoboziibigobabujijwekuvugaIKINYARWANDAigihecyosebarikuishuri>
- Verma, G.K. & Mallick, K. (1999). *Researching education: Perspectives and techniques*. London: Falmer press.
- Wall, D. (1997). Impact and washback in language testing. In C. Clapham & D. Corson (Eds.), *Encyclopedia of language and education (Vol. 7.)*. *Language testing and assessment* (pp. 291–302). Dordrecht: Kluwer Academic.
- Wang, C., Kim, D., Bong, M. & Ahn, H.S. (2013). Examining measurement properties of an English self-efficacy scale for English language learners in Korea. *International Journal of Educational Research*, 59 (2013), 24–34.
- Warner, G. (2016). *It's the no. 1 country for women in politics — but not in daily life*. Retrieved on 10 July 2017 from

<http://www.npr.org/sections/goatsandsoda/2016/07/29/487360094/invisibilia-no-one-thought-this-all-womans-debate-team-could-crush-it>

- Webb, N. (2002). Assessment literacy in standards-based education setting. *A paper presented at the annual meeting of the American Educational Research Association*, New Orleans, Louisiana. Retrieved on 18 August 2017 from <http://facstaff.wcer.wisc.edu/normw/AERA%202002/Assessment%20literacy%20NLW%20Final%2032602.pdf>
- Weir, C.J. (2005). *Language testing and validation: An evidence-based approach*. New York: Palgrave Macmillan.
- Wellington, J. (2000). *Educational research: Contemporary issues and practical approaches*. London: Continuum.
- Wellington, J., Szczerbinski, M., (2007). *Research methods for the social sciences*. (3rd ed.). London: Continuum International Publishing Group.
- White, J., & Lightbown, P. M. (1984). Asking and answering in ESL classes. *Canadian Modern Language Review*, 40, 228-244.
- White, P. (2009). *Developing research questions*. New York: Palgrave Macmillan
- William, D., Lee, C., Harrison, C. & Black, P. (2004). Teachers developing assessment for learning: impact on student achievement, *Assessment in Education: Principles, Policy & Practice*, 11(1), 49-65, DOI: 10.1080/0969594042000208994
- Williams, T. (2015). *Great expectations: Aspiration, uncertainty and schooling in Rwanda*. CfBT Education Trust.
- Wilson, E. & Fox, A. (2013). Data collection. In E. Wilson (Ed.), *School-based research: A guide for education students* (pp. 102-126). London: Sage Publications Ltd.
- Windschitl, M. (1999). The challenges of sustaining a constructivist culture. *Phi Delta Kappan*, 751-755.
- WNCP -Western and Northern Canadian Protocol- (2006). *Rethinking classroom assessment with purpose in mind*. Retrieved on 10 May, 2015 from <http://www.edu.gov.mb.ca/k12/assess/wncp>

- World Bank (2014). *Rwanda: National education profile*. Retrieved on 13 August 2016 from <https://www.epdc.org/education-data-research/rwanda-national-education-profile>
- World Bank (2017). *Rwanda: Overview*. Retrieved on 02 February 2018 from <http://www.worldbank.org/en/country/rwanda/overview>
- Yang, T.L (2008). Factors affecting efl teachers' use of multiple classroom assessment practices with young language learners. *English Teaching & Learning*, 32(4), 85-123.
- Zhang, Z. & Burry-Stock, J.A. (2003). Classroom assessment practices and teachers' self-perceived assessment skills. *Applied Measurement in Education*, 16(4), 323–342.
- Zimbicki, D. (2007). *Examining the effects of alternative assessment on student motivation and self-efficacy* (Preview). Walden University, ProQuest Dissertations Publishing. Retrieved on 13 March 2016 from <https://search.proquest.com/docview/304761711/previewPDF/25E2C4CA4C8A4718PQ/1?accountid=15181>
- Zimmerman, B. J. (1995). Self-efficacy and educational development. In A. Bandura (Ed.), *Self-efficacy in changing societies*. (pp. 202–231), Cambridge: Cambridge University Press.
- Zimmerman, J.B. & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment, *American Educational Research Journal*, 31(4), 845-862.
- Zimmerman, B.J., Bandura, A. & Martinez-Pons, M. (1992). Self-motivation and academic attainment: The role of self-efficacy beliefs and personal goals setting. *American Educational Research Journal*. 29(3), 663-676.