Self-assessment in EFL Speaking Classroom and its Effect on Achievement, Self-Regulated Learning, and Critical Thinking: Students' Voices from Saudi Arabia.

By

Bayan Fahad Alghanmi

A Thesis Submitted in Fulfillment of the Requirements for the Degree of Doctor of Philosophy (PhD)

School of Education
College of Social Sciences
University of Glasgow

December 2022

Dedication

I dedicate this thesis to Fahad and Baheya, my parents, for their endless love, support, and encouragement, my beloved husband Bassam, who has been a source of strength, support, and motivation throughout this journey, my beautiful children Mohammed and Maya, who arrived in this world during this journey and have been a source of motivation, my brother and sisters for always being there for me, and my in-laws, who are my second family, for their support and encouragement.

Abstract

In light of the growing need to enhance the quality of education to overcome social and economic issues, assessment systems and curriculum have undergone significant modifications and reforms in many countries. Saudi Arabia is no exception. The literature suggests that innovative approaches to assessment, such as self-assessment, have the potential to promote lifelong skills, empower learners, and enhance learning. Nonetheless, traditional assessment practices continue to dominate in Saudi Arabia, particularly in higher education English language courses. Review and reframing of assessment approaches are, therefore, necessary in Saudi Arabia to improve the quality of learning and to develop learners' lifelong skills, including self-regulated learning skills and critical thinking skills. Despite the growing interest in self-assessment as a practical instructional strategy that draws on formative assessment to promote self-regulated learning and critical thinking, relatively few studies have addressed this topic in English language courses in higher education, and none have addressed it in the context of Saudi Arabia. The evidence regarding the impact of self-assessment on the quality of learning and the empowerment of learners may help to guide the Saudi education reform.

Nonetheless, traditional assessment practices continue to dominate in Saudi Arabia, particularly in higher education English language courses. Therefore, reviewing and reframing of assessment approaches, specifically to improve the quality of learning, are necessary in Saudi Arabia to develop learners' lifelong skills, including self-regulated learning skills and critical thinking skills. Recently, self-assessment has emerged as a practical instructional strategy that draws on formative assessment to promote self-regulated learning. However, most research on formative assessment and self-regulated learning has portrayed results related to self-assessment as generalisable, despite the need for research across various educational contexts.

The aim of this research is to explore in depth the participants' perceptions and experience of self-assessment in speaking classrooms and the impact of self-assessment on learners' self-regulatory skills, critical thinking, and speaking language performance within the EFL context. This study also examines the relationship between learners' self-regulated learning and their critical thinking skills. Pre- and post-tests were conducted with 27 EFL students before and after a self-assessment intervention. In addition, a self-assessment proforma, audio recording, and semi-structured interviews were collected and conducted with 10 of

the 27 students. All these tools played an essential role in investigating the participants' perceptions and experience of self-assessment and its impact.

Overall, the participants in this study displayed favourable attitudes towards self-assessment. The findings indicate that a variety of factors influenced the learners' perspectives, including learners' prior experience with traditional speaking assessment, learners' motivation and willingness to self-assess, learners' awareness of assessment criteria, and learners' perceptions and experiences of feedback. The findings also reveal the positive impact of self-assessment on learners' self-regulated learning skills, critical thinking skills, and achievement, as well as a positive medium-strength relationship between learners' self-regulated learning skills and critical thinking skills.

The study concludes with recommendations for educational policy-makers who are aiming to establish practices that support and empower learners. For example, the study encourages the use and adaptation of the self-assessment proforma in the English language as a reliable scaffolding method of assessment that can foster deep learning and self-regulated learning.

Table of content

List of tables

List of figures

List of Appendices

Acknowledgments

Firstly, I would like to express my gratitude to Allah for giving me the strength to complete this thesis.

I also offer special thanks and gratitude to my first and second supervisors Doctor Georgina Wardle and Professor Louise Hayward for their guidance, support, assistance, and commitment throughout this journey. Thank you for always pointing me in the right direction towards achieving my dream. I learned a lot from your valuable insights, and I am forever grateful.

My love and gratitude for all of my loved ones is indescribable, particularly my parents, my husband, children, brother and sisters, and second family (my in-laws). Thank you for everything and for always keeping me in your prayers. Without the inspiration and support that you provided, I would not have been able accomplished this goal.

Special gratitude also goes to the government of the Kingdom of Saudi Arabia, as represented in the United Kingdom via the Ministry of Education and the Cultural Attaché and the University of Jeddah (UJ) represented by the English language institute (ELI) for granting me the scholarship and supporting me throughout this journey.

Author's declaration

I declare that, except where explicit reference is made to the contribution of others, that this dissertation is the result of my own work and has not been submitted for any other degree at the University of Glasgow or any other institution.

Bayan Fahad M Alghanmi

_____

Abbreviations

AaL	Assessment as Learning
AfL	Assessment for Learning
ALM	Audio-Lingual Method
AoL	Assessment of Learning
CLT	communicative language teaching
CT	Critical thinking
CTA	Watson-Glaser II Critical Thinking Appraisal
EFL	English as a Forging Language
ELT	English Language Teaching
FA	Formative Assessment
GMT	Grammar Translation Method
HE	Higher Education
HEI	Higher Education Institute
KSA	Kingdom of Saudi Arabia
IPA	Interpretative Phenomenological Analysis
LLL	Life Long Learning
MoE	Ministry of Education
SA	Summative Assessment
SR	Self-regulation
SRL	Self-regulated learning
SRL-SRS	Self-Regulation of Learning Self-Report Scale
UK	United Kingdom of Britain

Chapter 1 : Introduction

1.1. Introduction

As one of the three pillars in the educational system, namely curriculum, pedagogy, and assessment, assessment plays a significant role in the development of education, including in processes such as teaching, learning, and decision-making (Coombs et al., 2018). Rather than merely an extra step at the end of the teaching and learning process, assessment is an integral part of the process according to Brown and Glasner (1999). In fact, assessment can influence the learning process in beneficial ways to improve students' learning. It can involve the students in meaningful learning and help them to understand how well they are progressing and what else they must do to advance their learning (Brown & Glasner, 1999). It can, thus, inform future learning and prepare students for lifelong learning (LLL). However, to accomplish these goals, assessment needs adjustment and reform (Boud & Falchikov, 2007). Considering the rising need to improve the quality of education to address and overcome social and economic problems, many countries have come to view assessment as playing an increasingly important strategic role in education policy formation (Livingston & Hutchinson, 2017). Consequently, many countries around the world have attempted to make significant modifications and reforms in their assessment systems and curricula over the past two decades, and Saudi Arabia is no exception.

In 2006, assessment emerged as the practice that required the most reformation and improvement in a report on quality assessment of education programmes by the United Nations Development Programme Regional Bureau for Arab States (UNDP/RBAS, 2006) in 23 Arab higher education institutions (HEI), including Saudi Arabian HEI. Assessment in Saudi Arabia is mostly seen as a means of grading students on the basis of the information they have acquired by the end of the academic year or course (Darandari & Murphy, 2013). This view of assessment leads to teacher-centred classrooms, in which the teacher is the source of information while students tend to be passive receivers of knowledge (Heywood, 2000). The teacher-centred approach can hinder the development of lifelong skills that could otherwise empower learners and improve their learning (Heywood, 2000). Thus, a transition from assessment as judgement is needed to support learning and help students become self-regulated learners and develop lifelong skills like critical thinking. This transition, however, is not free from challenge due to the top-down

assessment policy, which is the dominant state model in the Kingdom of Saudi Arabia (KSA), and the strong influence of traditional views of teaching, learning, and assessment (Do Quyen & Khairani, 2017).

Over the last decade, higher education (HE) in Saudi Arabia has begun to question its dependence on teacher-centred classroom, including traditional assessment methods (Darandari & Murphy, 2013). However, despite attempts to shift from the teacher-centred and traditional assessment approaches to a new assessment culture, traditional assessment practices still dominate in Saudi Arabia, especially in the English language courses in HE (Almossa, 2021). Hence, assessment approaches should be reconsidered and reframed specifically in English language courses in higher education to boost the quality of learning in Saudi Arabia and develop learners' lifelong skills.

1.2. The motivation for the study

Personal and cultural reasons have motivated me to investigate the impact of self-assessment in speaking classrooms on English as a foreign language (EFL) learners' self-regulated learning and critical thinking in the Saudi Arabian EFL context.

This research was first inspired by my personal desire to see self-assessment, self-regulated learning, and critical thinking extensively acknowledged and practised in Saudi higher education. My initial exposure to the concepts of self-regulated learning and critical thinking occurred during my master's (MA) in TESOL, which I took in Saudi Arabia. I was immediately intrigued by the concepts and began analysing and contrasting my own learning behaviour during my MA with my past learning behaviour during my Bachelor's (BA) in English Language, which I also took in Saudi Arabia. During my MA, I experienced self-regulated learning, in contrast to my BA, in which students were encouraged to be passive receivers in class, where the focus was always on the instructor as the source of information. Experiencing self-regulated learning made me feel that undergraduates in Saudi Arabia should have access to these opportunities for learning to develop skills for the future, or LLL skills. In addition, from my own teaching experience in a higher education institute (HEI) in KSA, I have noticed that the majority of students adopt a passive role instead of engaging and trying to take an active role in class. My students tended to wait for information and directions from the teacher. This teaching experience, in turn, further inspired me to investigate and gain initial insights into how Saudi learners perceive the notion of self-regulated learning.

Consequently, I asked the students in my English language classes about the idea of self-regulated learning to gain a sense of their broad thoughts on the concept. Self-regulated learning seemed to be a new concept to the students, yet there was a favourable response in terms of its potential benefits, and some students asked about how to implement and develop this approach to learning. However, research in EFL has suggested that effective learning in EFL requires not only self-regulated learning but also critical thinking (see, for example, Ghanizadeh (2017); Ghanizadeh and Mirzaee (2012); Kamgar and Jadidi (2016)). I decided then to pursue an investigation into the concept in higher education in Saudi Arabia as a contribution to enhancing the standard of language learning in the Saudi HEIs and assisting the learners in developing LLL skills.

As a person who has lived, studied, and worked in educational settings in Saudi Arabia, I have experienced the obstacles of EFL learning as a student and observed them as a teacher. Assessment in Saudi Arabia is mostly seen as a means of grading students on the basis of the information they have acquired by the end of the academic year or course (Darandari & Murphy, 2013). Adherence to traditional teaching and assessment practices continues in HE particularly in the English language courses, notwithstanding efforts to shift from the traditional teaching and assessment culture (Almossa, 2021). This emphasises the need to shift away from traditional practices of assessment to support enhancing the quality of learning in English language courses in higher education in KSA.

In recent years, Panadero et al. (2017) and Brown and Harris (2014) proposed and tested self-assessment as a practical instructional strategy that drew on formative assessment to support self-regulated learning. Specifically, the researchers tested its effectiveness in promoting self-regulated learning and learners' empowerment, and improving learning. However, what is effective in a context where English is the native language of students may not necessarily be effective in a context where English is taught and learnt as a foreign language. Panadero et al. (2018), in a review paper, noted that most research on formative assessment and self-regulated learning has presented the findings as though they are generally applicable; thus, more attention should be given to researching the relationship across various educational contexts (Panadero et al., 2018).

Therefore, prior to implementing any changes or reforms, there is a requirement to study, identify, and discuss information and practices related to the impact and experience of self-assessment in the Saudi context. More precisely, the impact of self-assessment on self-regulated learning, critical thinking skills, and achievement must be examined, as that

impact may be unique and perceived differently in the Saudi Arabian context. Additionally, learners' perceptions and experience within the learning process are fundamental to the success or failure of any attempts to reform assessment (Kamgar & Jadidi, 2016). Several researchers have highlighted the importance of investigating assessment practices from language learners' perspectives (Alhamami, 2019; Borg & Edmett, 2019; Kalra et al., 2017; Tong, 2011). Therefore, it is also significant to explore EFL learners' perceptions and experience of self-assessment and its impact on learners' self-regulated learning, critical thinking, and speaking language achievement. Notably, no research has investigated the effect of self-assessment on learners' self-regulated learning and critical thinking in Saudi Arabia, particularly in English language courses in HE. I decided to pursue an investigation into the influence of self-assessment on learners' self-regulated learning, critical thinking, and language achievement in HE in Saudi Arabia to contribute to the provision of an evidence base that could be used to help enhance the language learning quality in higher education and assist undergraduates in developing LLL skills.

1.3. Significance of the study and research gap

This research investigates the impact of self-assessment on learners' self-regulated learning and critical thinking in EFL speaking classes. Relatively few studies have addressed this topic from the learners' perspective, despite wide recognition of the critical role of learners' perceptions. Of the studies that have addressed this topic (see, for example, Kahrizi et al. (2014) and Ghanizadeh and Mirzaee (2012)), none have addressed it in the context of Saudi Arabia. The topic of self-regulated learning and critical thinking at the higher education level is critical to the Saudi government's 2030 vision and its goal of developing a knowledge economy. A knowledge economy emerges with increasing usage of knowledge-based outputs obtained via intellectual capabilities; such knowledge-intensive tasks demand independent thinkers who are capable of creating and solving problems (Powell & Snellman, 2004). Despite the topic's relevance in the Saudi context, this research gap remains. These subjects, as well as the relationship between self-regulated learning and critical thinking, have not been thoroughly examined in higher education and in an EFL context.

In this research, which is situated within an EFL context, there is a need to investigate and comprehend learners' perspectives of self-assessment, as well as to comprehend how it supported or impeded student learning and development. The hope is that this research

increases interest in the implementation of innovative assessment in EFL classes in KSA. The findings of this research may be used to encourage language teachers in the Saudi context to consider implementing self-assessment in the English language classroom as a means to empower learners through promoting self-regulated learning and critical thinking and thus improve the quality of learning. The findings of this research may also provide stakeholders and teachers in Saudi higher education with a framework of knowledge on how students perceive the use and impact of self-assessment and stimulate questions, opportunities, and possibilities for its applicability. Such a framework could help educators to overcome difficulties in self-assessment implementation with English language learners who lack experience in self-assessment.

1.4. The purpose of the study

The aim of this research is to explore in depth the participants' perspectives and experience of self-assessment in speaking classrooms and the impact of self-assessment on learners' self-regulatory skills, critical thinking, and speaking language performance within the EFL context. This study also examines the relationship between learners' self-regulated learning and their critical thinking. The participants were EFL learners in their preparatory year at a higher education institution where they are learning English.

1.5. The research questions

The following questions were formulated to address the aim of this research:

1) How do EFL students perceive the implementation of self-assessment in speaking classes, especially in relation to their speaking language achievement?
2) What influence does self-assessment in speaking have on EFL students regarding:
    (a) their self-regulated learning in English language speaking?
    (b) their critical thinking skills?
3) What is the association between students' perceptions regarding the influence of self-assessment in speaking on their self-regulated learning and their critical thinking?

1.6. Thesis structure

In the first chapter *(Introduction)*, the rationale for the study is established and the purposes of this research are defined. This chapter has also outlined the study's

significance and the gap and research questions that the study aimed to address. The Saudi Arabian context, which serves as the context for this study, is thoroughly described in the second chapter of this thesis, *Context of the Study: Saudi Arabia*. This second chapter details the Saudi education system and discusses the system's objectives in relation to this research. A comprehensive understanding of the significance of the English language and its function in the Saudi setting and education system is also provided in this chapter. Finally, the chapter provides a synopsis of the materials, objectives, approaches, and assessment used in English language teaching in Saudi Arabia.

The third chapter, namely *Literature review,* offers a thorough review of the relevant literature to this study. This chapter first discusses the concepts of self-regulated learning and critical thinking and considers the connection between the two. The chapter then presents and discusses the key concepts, the conceptualisation of assessment over time, and self-assessment, including its definitions and its contextualisation in the constructivist and sociocultural paradigms. Finally, the chapter explores the potential of self-assessment as a means of enhancing critical thinking and self-regulated learning in an EFL context. In the fourth chapter, the methodology employed in gathering the data and the data analysis in the study is outlined. This chapter presents the rationale for using a mixed methods approach and the advantages associated with that approach and discusses the tools used to collect the data and the rationale for the study sample. Additionally, it details the steps taken to ensure that the study upheld all applicable ethical standards. Finally, the chapter considers validity and reliability in detail. The next two chapters present findings from the analysis.

The fifth chapter in this thesis, which contains the *Quantitative data analysis,* outlines the methods and processes used in the analysis of the pre- and post-tests and presents the findings of both tests. In turn, the sixth chapter, which contains *Quantitative and qualitative data analysis,* outlines the methods and processes used in the analysis of learners' audio recordings and self-assessment proforma and presents its findings. This chapter also outlines the methods and processes used in the analysis of the semi-structured interviews with the EFL learners, and presents the main themes drawn from the interviews.

In order to address the research questions, the seventh chapter (*Discussion)* discusses the findings of this research, presented in the fifth and sixth chapters, together to situate the findings in the context of wider literature. The last chapter in this thesis provides a conclusion for the study by summarising and reflecting on the research key findings. It also identifies possible implications for educational policy and higher education institutes and

practice. It identifies the limitations of the study and offers several recommendations for further studies.

Chapter 2 : Context of the Study (Saudi Arabia)

2.1. Introduction

The chapter outlines the background and context of the Kingdom of Saudi Arabia (KSA), within which this research is situated. It reviews the English language teaching history in Saudi Arabia by tracking the ways in which, over time, thinking has changed about the teaching of English. The chapter then examines how the English language was incorporated into Saudi education over time. The chapter continues by describing the education system in Saudi Arabia and analyses education policies and objectives in the Saudi context that are of particular relevance to this study. Finally, the chapter reflects on English language materials, teaching methods, and assessment used in the Saudi EFL context.

2.2. Profile of the Kingdom of Saudi Arabia

The Kingdom of Saudi Arabia (KSA) is a Muslim Arab country founded by King Abdul-Aziz Al-Saud in 1902 (Nonneman, 2001). It is situated in the south-west region of Asia and is thus geographically at a crossroads between Africa, Europe, and Asia. It is split into 13 administrative divisions, or regions, including Riyadh, the capital, in the central area; Dammam and Jeddah, the two principal ports; Makkah and Medina, the Islamic holy cities; and Abha, in the southern region. The Saudi Arabian economy was modest till World War II, relying on subsistence farming, limited commerce, pearl fishing, camel exports, and the income earned from people paying to undertake pilgrimage to one of the world's major religious sites. Although oil was discovered in the eastern region in 1938, it was not widely utilised until the Arab American Company was founded in 1946 (Al-Sadan, 2000). Since then, Saudi wealth has influenced every part of society, altering societal values, increasing educational and healthcare possibilities, and enhancing the standard of living (Al-Sadan, 2000). Moreover, with the two holy cities for all Muslims, Makkah and Medina, the KSA is widely considered to be the heartland of Islam. The Arabic language, which is the holy Quran's and Islam's sacred language, is the official language in Saudi Arabia, and all elements of society, including the education system, are governed by it (Oyaid, 2009).

2.3. Saudi education system

The education system in Saudi Arabia, which includes general education and higher education, focuses on the teaching of Islam, has a centralised system of supervision and educational assistance, receives governmental funding (thus, education is free at all levels in the KSA), and has a general policy of gender segregation (Smith & Abouammoh, 2013). The Saudi educational system follows a top-down policy in which the Ministry of Education (MoE) is permitted by the Saudi government to make all decisions on education, including curriculum, instructional practices, and educational materials. (Nunan, 1989). In general education, there are 12 years of education: six years of primary or elementary education (ages 6–12 years old), three years of intermediate education (ages 13–15), and three years of secondary education (ages 16–18). Higher education in the Kingdom of Saudi Arabia is a relatively new concept, with the first university being founded in 1957. The Higher Committee for Educational Policy in the MoE, which is accountable for both general and higher education, defines the objectives of Saudi education based on the teachings of Islam as follows:

1. To provide learners with the vital information and abilities to become valuable members of society.
2. To develop learners' sense of cultural, economic, and social issues and prepare them to contribute to these problems' solutions.
3. To reinforce individuals' dignity and give individuals equal opportunities to develop their skills to participate in the development of the country.
4. To motivate and develop scientific research and thinking, strengthen the ability to observe and reflect, and enlighten students about the signs and miracles of God in the universe all in order to enable the students to play an effective role in building social life and properly directing it.
5. To teach learners about great achievements in literature, science, and other fields, revealing that scientific progress results from the efforts of all mankind.
6. To develop mathematical thinking, arithmetical skills, reading skills, and reading habits and to train students in the use of the language of figures and its applications in the scientific field.
7. To teach students to express themselves correctly in speech and in writing.
8. To teach students at least one foreign language so they can benefit from it.
9. To view each student as an individual and direct them and help them to grow in a way best suited to their abilities.
10. To give students the opportunity to do manual work and gain experience in laboratories and in building and agricultural work.
11. To study the scientific principles of various activities in order to encourage progress and innovation in mechanical production.

(MoE, 1970)

The policy context for this study might seem favourable, as it could be argued that most of

the objectives in the Educational Policy highlight the importance of self-assessment, self-regulated learning, and critical thinking in learning and require students to improve and develop these skills in order to achieve these objectives. Promoting self-assessment, self-regulated learning, and critical thinking could offer opportunities to achieve these objectives in learning.

For example, three objectives concern encouraging learning that could help students take on active roles in their society (objectives 1, 2 and 3). Promoting self-assessment, self-regulated learning, and critical thinking have the potential to help students be active members of their society, as well as actors capable of solving social and cultural problems. Namely, self-assessment supports learners to participate actively in the assessment process, prompting them to think more deeply and develop crucial cognitive skills such as critical thinking, collaboration, decision-making, self-monitoring and regulation, and problem-solving; self-assessment pushes learners to criticise constructively and suggests improvements, reflect, and make sensible judgements (Sung et al., 2005). The role of self-assessment in the learning process is discussed later in Chapter 3 (Section 3.3, p. 38). Thus, the impact of self-assessment, self-regulated learning, and critical thinking on learning could enable students to become better learners. These factors could lead students to become more skilful and able to become active members in their societies. Such graduates could participate in resolving social issues instead of only acquiring knowledge; indeed, knowledge alone is insufficient to enable learners to participate and be involved in their societies. Moreover, the educational objectives of Saudi Arabia place importance on the individual as a part of the society. Objective 9 emphasises the role and the growth of the individual and highlights the importance of providing guidance based on an individual's unique abilities. Introducing self-assessment and self-regulated learning in the Saudi education system would encourage independence and enrich the view of individuality. Self-assessment and self-regulated learning enable learners to reflect on their learning and to connect with their existing knowledge and, in turn, may prompt students to learn by themselves. Students could thus attain several objectives of the education system in Saudi Arabia through employing strategies of self-assessment. This thesis returns to these three concepts (self-assessment, self-regulated learning, and critical thinking) in Chapter 3.

Another of the Saudi's Educational Policy's objectives, Objective 8, highlights the importance of introducing at least one foreign language into the curriculum. In this case, the English language has been chosen by the Saudi government. The ability to use the English language has grown in significance, predominantly in business and education, with

the spread of globalisation (Nouraldeen & Elyas, 2014). Additionally, the necessity for the English language in global communication has contributed to the spread of language (Nouraldeen & Elyas, 2014). Kachru et al. (2009) identified three categories of English that are frequently referred to as the three concentric circles of English; together, these are sometimes referred to as the global English circle model. This model includes the inner circle, the outer circle, and the expanding circle (Figure 2-1). The inner circle denotes nations where English is spoken as a first language, such as the United States, the United Kingdom, New Zealand, Australia, and Canada, and indicates the origins of the English language (Kachru et al., 2009). The outer circle denotes nations where the English language was introduced through colonialism; English is used as a second language alongside the country's primary language, even in official communications in these countries. Outer circle nations include Malaysia, India, and several African countries (Kachru et al., 2009). The last circle, the expanding circle, includes nations where English as a foreign language (EFL) is spoken but not widely utilised and has no significant role. The expanding circle refers to the remaining countries around the world, such as China, Korea, and Saudi Arabia (Kachru et al., 2009). Saudi Arabia recognises the importance of the English language in international communication. Hence, English is the only foreign language introduced as a compulsory subject in the Saudi education system, and it is currently taught in primary (from grade two, which is around 7–8 years old), intermediate and secondary schools, and higher education.

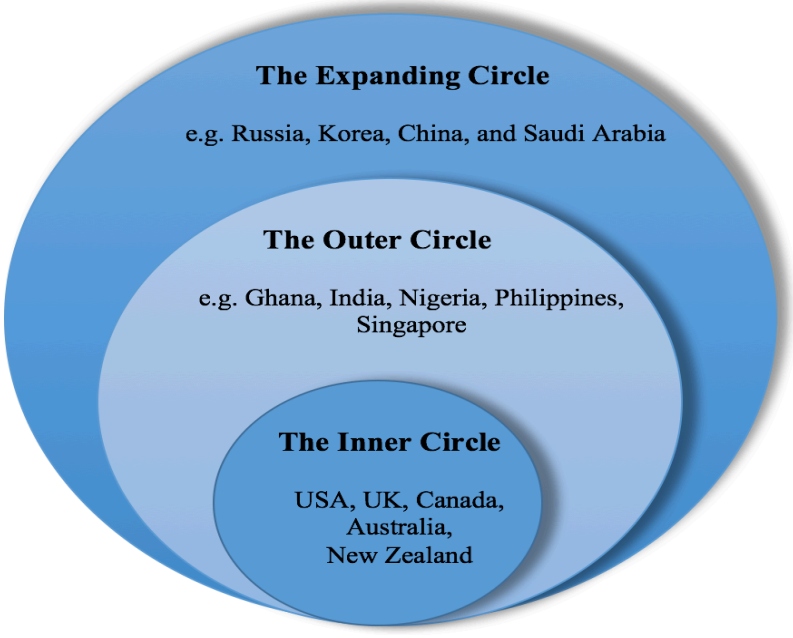


Figure 2-1 The global English circle model (adapted from Kachru et al., 2009)

Regardless, as Al-Issa (2009), Qahtani (2010), and Alharbi and Madhesh (2018) crucially observed, the policies and purposes of the Saudi education system were developed in 1970; and since then, few changes have emerged. The lack of change remains the basis of much criticism despite the presence of critical thinking, self-regulated learning, and self-assessment in the Educational Policy objectives. It is no longer logical or rational to apply the same policy and purposes with the same words that were designed for a different era; this language, at the very least, needs refining and reforming in accordance with the present context and requirements (Qahtani, 2010). Al-Issa (2009) argued that cultural and societal changes lead to alterations in educational systems and policies. Al-Issa (2009) proposed that a new educational policy should embrace change through encouraging discovery, searching, and inquiry and the strengthening of mental abilities among the new generations; such areas of growth subsequently enhance students' cognitive and metacognitive skills and, ultimately, their prospective productivity, while providing students with the skills and knowledge to succeed in their individual and social lives.

However, to date, relatively little research has been undertaken that has focused specifically on the Saudi education system, especially related to policies, strategies, and initiatives for assessment that define assessment objectives, strategies, and implementation. This lack of research is particularly concerning in light of the release of the National Transformation Programme and Saudi Vision 2030 in 2016. This vision aims to reduce Saudi Arabia's reliance on oil and to ensure the Kingdom is a thriving society, a prosperous economy, and an ambitious nation (Saudi-Vision-2030, 2016). The Saudi Vision is dedicated to developing a prosperous country in which all residents may accomplish their aspirations. This vision is also supported by the MoE and its educational policy; however, to achieve this vision, the country should reform its education system with a focus on connecting education objectives with economic development to produce a skilled and educated workforce that can advance the nation economically.

Significantly, as Alharbi and Madhesh (2018) noted, the Educational Policy is the main policy followed by all schools and higher education institutions in the KSA, and the English language is the only foreign language that has been introduced as a compulsory subject in the Saudi education system, a fact also emphasised by Gaffas (2019). Therefore, prior to taking any action, it is necessary to investigate what would work best in the Saudi Arabian context. The following section addresses the historical background of the English language in Saudi Arabia.

2.4. English language in Saudi Arabia

As identified in Section 2.2 of this chapter, the oil discovery in the KSA in the 20th century, affected every aspect of life and changed social values. It resulted in significant economic expansion, which in turn resulted in the hiring of a large number of workers from foreign countries including the Philippines, Indonesia, India, and Pakistan (Al-Sadan, 2000). These workers came to work in different locations around the country such as governmental and private companies, banks, and hospitals. The resulting need to interact with the international community, as well as with international workers within Saudi Arabia, encouraged the Saudi people to learn the English language due to its importance in the international community (Al-Motairi, 2005). Accordingly, English Language Teaching (ELT) and learning in the Saudi area began to increase (Javid et al., 2012). Four reasons for teaching English in Saudi schools were highlighted by Al-Hajailan (2003). First, English is regarded as the first global language spoken in both the East and the West. Second, most published resources, including international research, are written in English. Third, English is the international language of business and commerce. Lastly, it is the United Nations' official language, and it is spoken in most nations for a variety of reasons including trade, education, and tourism. Consequently, Al-Seghayer (2011) highlighted that, following the establishment of the Directorate of Education in 1923, the government of Saudi Arabia introduced English as a compulsory intermediate education course in 1927.

Between 1923 and 1953, the Directorate of Education introduced a variety of educational goals pertaining to English Language Teaching (ELT). One example was the introduction of ELT into the curriculum in Saudi Arabia in 1948 through the establishment of an industrial school, or a training school that teaches students the skills they need to work in industry, with a curriculum adopted from Egypt and Egyptian teaching staff. The industrial school taught the usual theoretical subjects including science and mathematics, the practical subjects appropriate to a vocational school, and the English language (Al-Subahi, 1988; Lebeeb, 1993). Nonetheless, the syllabus was deemed to be unsuitable for the needs of Saudi students, since it no longer matched the criteria of the Saudi education system and did not contribute to its primary goals (Al-Subahi, 1988). The 1960s witnessed the appearance of the first comprehensive English curriculum in both intermediate and secondary schools. This curriculum continued until the 1970s, as reported by Al-Seghayer (2005), when policy changes led to eight hours of English teaching per week being reduced to six. However, in the 1980s, a new English curriculum known as Saudi Arabian Schools

English (SASE) was launched and replaced the previous programme, which was deemed not appropriate for learners' linguistic requirements. The new English programme set out to reach the goals of the Saudi students' requirements (Al-Subahi, 1988).

Another initiative in English Language Teaching (ELT) was launched by native-speaking (NS) specialists in the 1990s and implemented under the formal supervision of King Fahd University of Petroleum and Minerals (KFUPM), the university where the NS specialists worked (Al-Seghayer, 2011). The MoE initiated this programme by releasing a series of intermediate and secondary school textbooks entitled *English for Saudi Arabia* (Al-Seghayer, 2011). Nonetheless, Almulhim (2001) argued that these textbooks may not have been suitable at that time in the Saudi context; thus, have been found unsuitable for the learners' requirements. Indeed, the textbooks have been criticised for neglecting certain aspects of Saudi society and culture. Until then, English had been taught in intermediate and secondary schools; however, since evidence reveals that the majority of high school and college graduates in Saudi Arabia had low levels in English language proficiency (Al-Seghayer, 2005), the Higher Committee on education policy in the MOE designed a new programme. The new programme aimed to introduce English into primary schools initially at the sixth grade (2004) and then later in the fourth grade (2008) as a further attempt to improve English learning (Elyas, 2008).

Considering the significance of the English language in the international community (Al-Motairi, 2005), English has been taught at all levels of higher education in the Kingdom of Saudi Arabia, including universities, military academies, and technical and vocational schools. When King Fahad University was formed in 1975, it was the first to use English as a medium of instruction, as was King Abdullah University of Science and Technology (KAUST) on its foundation in 2009. Moreover, in 2004 a major policy shift took place when all universities in the KSA were required to introduce a preparatory year programme that included a significant time allocation to ELT covering all aspects of English including listening, speaking, reading, writing, grammar, and vocabulary (Al-Motairi, 2005). This programme was gradually introduced in the universities over the next five years. Currently, one of the requirements for acceptance into university departments in Saudi Arabia such as engineering and medicine is an acceptable level of English. Additionally, English is the language of instruction for a variety of courses in higher education in KSA, including engineering, computer science, and medicine. The following section addresses the objectives of teaching English as a Foreign Language (EFL) in the KSA and the methods and materials used.

2.5. English language teaching objectives, materials, and approaches in Saudi Arabia

Enhancing students' language proficiency is considered as the main goal of teaching EFL in the KSA. Al-Motairi (2005) argues that EFL in the KSA has been taught with the aim of increasing students' language proficiency in order to ensure students can access and understand international references, since many publications in different fields are written only in English. Thus, intellectual, personal, and professional development depends on students' English language ability. For example, the new English programme at one Saudi higher education institution (HEI) aims to build students' English skills in key areas, both academically and personally through building students' English proficiency in the four language skills. Further, the programme aims to enhance these skills with linguistic and lexical competencies as well as to develop thinking skills, presentation skills, and sub-skills. Part of teaching English in institutions is also that developing students' cognitive skills is extremely important. Integrating strategies of self-assessment into EFL pedagogy therefore makes sense, as doing so could help enhance learners' competence in the language as well as their cognitive skills and ultimately make them better learners. The developments in English language teaching in the KSA, including the materials and teaching methods, were all attempts to achieve the goal of improving the quality of EFL teaching in the country.

In an attempt to improve learners' language proficiency in the KSA, schools and universities began to use textbooks and materials developed in countries where English is the first language (Elyas, 2008). It is important to note that policy in the KSA does not allow teachers to adapt textbooks to the needs of their learners or their learners' situations; without any deviation, teachers must adhere to the textbooks (GDA, 2013). This inflexibility, according to Shah et al. (2013), could negatively impact the teaching and learning process. It de-professionalises teachers and limits them to using only the traditional teaching methods used in the textbooks, methods that tend to produce only surface learning and passive acceptance of ideas and information. Shah et al. (2013) discussed that additional factors are thought to influence English instructional methods along with the ministry of education's constraints on the content and methods of teaching English in Saudi Arabia. EFL teachers in Saudi Arabia are all given the same syllabus, along with standards and deadlines that they must adhere to. Additionally, instructors' perceptions of language learning, as procedural memorising of grammar rules and vocabulary, is another explanation for the dominance of traditional teaching approaches;

teachers prefer to organise classes and content around a presentation of the information that students must remember and acquire. Additionally, the perceptions of the instructors regarding interaction with students only includes learners responding to questions posed by the instructor. Teachers' perceptions and attitudes thus contribute to the development of an exam-oriented culture.

Traditional teaching approaches to language teaching and learning, such as the grammar translation method (GTM), the direct method, and the audio-lingual method (ALM), have been commonly used in language teaching in the KSA (Al-Awaid, 2018; Al-Seghayer, 2005; Al-Seghayer, 2011, 2015; Khan, 2011; Mahmoud, 2012). These methods have also been used around the world in language teaching. Today, though, these methods have been rejected in most countries, as they have been shown to be ineffective in enhancing English language proficiency and communicative competence (Griffiths & Parr, 2001). However, in the KSA, these traditional teaching approaches are still in use. Predominantly, teacher-led presentations and explanations of new language elements are used, with limited possibilities for student involvement. In the KSA, GTM tends to be used the most in EFL classrooms (Al-Seghayer, 2005; Al-Seghayer, 2011, 2015; Khan, 2011). Widespread use of the GTM remains, despite the issues with its use. GTM has been deemed to be a traditional, teacher-centred approach. It is underpinned, first and foremost, by conducting an in-depth study and analysis of grammar rules and then applying the gained knowledge into practice through translation assignments to practise translating to and from the target language (Richards & Rodgers, 2001, 2014). In other words, GTM mainly concentrates on the teacher's presentation skills and translation. According to Richards and Rodgers (2001, 2014), GTM places a strong emphasis on comprehending the system of the foreign language (English in the Saudi context) instead of using the English language in communication. GTM concentrates on teaching rather than learning (Al Asmari, 2013). GTM's focus is limited to the application and mastery of linguistic patterns; it pays insufficient attention to the use of those patterns in communicative situations. Thus, rather than imparting communicative knowledge, GTM gives learners only formal and technical knowledge. Teaching approaches like those advocated by GTM tend to offer only limited opportunities to exercise self-regulated learning and critical thinking through self-assessment. Such approaches encourage learners to be passive receivers during classes and heavily dependent on their teachers rather than on themselves. Self-regulation or articulation of thinking during learning is thus extremely challenging for learners accustomed to such approaches.

The direct method is another approach to teaching and learning that is still widely used by Saudi EFL teachers (Al-Seghayer, 2005; Al-Seghayer, 2011, 2015; Khan, 2011). The direct method concentrates on vocabulary learning through presentations and then asking questions to check students' understanding (Richards & Rodgers, 2001, 2014). The direct method also tends to keep learners as passive receivers of knowledge from their teachers; it does not train learners to use or practise using the knowledge (Al-Seghayer, 2005). In addition, Zaid (1993) claimed that, by applying the direct approach, EFL language instructors sometimes implement cognitive code learning, which fosters the conscious selection of important forms of grammar throughout the learning process. Cognitive code learning heavily concentrates on correct grammar forms. Richards and Rodgers (2001, 2014) suggested that this method prioritises grammatical forms of the language over meaning and argued that this may cause problems for students at the early stages of language learning. To promote self-regulated learning and critical thinking in language learning instead, teachers need to shift towards methodologies that elicit meaningful interaction.

The audio lingual method is another approach to teaching and learning the English language in Saudi Arabia (Al-Seghayer, 2005; Al-Seghayer, 2011, 2015; Khan, 2011; Zaid, 1993). The ALM also focuses on structure and form rather than on meaning, and it emphasises memorisation methods (Richards and Rodgers, 2001, 2014). The ALM is dependent on teaching and learning foreign language through drilling and repetition of words and grammatical rules in language laboratories (Castagnaro, 2006). This dependence on such facilities is problematic given some schools lack the necessary facilities. Zaid (1993) argued that EFL instructors incorporate ALM into their classes in order to adhere to the established standards for English instruction. Zaid (1993) further argued that there is minimal overlap between the objectives of education and the utilised methods in EFL teaching and learning in the Saudi context. Namely, Zaid (1993) argued that the ALM, which is grounded on drill and repetition, would be ineffective in improving learners' cognitive abilities, which is a key objective of English instruction. Hence, the ALM is not suitable for promoting self-regulated learning and critical thinking in English language learning through self-assessment. It concentrates on repetition and memorisation, whereas self-assessment, self-regulated learning, and critical thinking centre on enabling learners to use their cognitive skills and enabling learners to make meaning of their learning.

All of the discussed teaching methods used in Saudi Arabia (the GTM, direct method, and ALM) put greater emphasis on grammatical forms and content instead of practising the

actual use of language in real-life situations. Accordingly, these methods are unlikely to help learners to become active learners and develop their intellectual abilities; consequently, these methods are unlikely to contribute to the achievement of the objective of raising the quality of English language instruction in Saudi Arabia. To achieve this goal, EFL teaching in the KSA needs to transform to learner-centred instead of teacher-centred. This shift may be best achieved through the adoption of alternative pedagogical approaches such as the communicative language teaching (CLT) approach, a student-centred approach that shifts the focus from teachers to students. Mahmoud (2012) clarified that, in the late 1990s and early 2000s in the Saudi HE sector's language instruction, the CLT approach began to emerge and develop. This period is thus a significant period in terms of the focus shifting from teacher-centred to learner-centred education.

Warschauer and Kern (2000) also observed that the traditional methods long practised in language teaching, such as the ALM and GTM, began to be transplanted by the CLT approach at the turn of the century. Hymes (1972), who originally developed the CLT approach, prioritised communicative use of the actual language being learnt over simple clarification of lexical, grammatical, and phonological sets of rules (Hiep, 2007). Interactions within the communicative classroom indeed improve language learning more than mere repetition of regular tasks (Consolo, 2006). CLT, as described by Richards and Rodgers (2001, 2014), Brown (2000), Richards (2006), and Ariza and Zainuddin (2002), consists of a set of norms and principles for language teaching that covers the objectives, classroom activities, and responsibilities of the instructor and learners through communication and interaction. It has been claimed by Brown (2000), Richards (2006), Hiep (2007), and Ariza and Zainuddin (2002) that communicative competence is to be regarded as the primary objective of language learning and deemed necessary for the effective use of the target language. Communicative competence, as outlined by Richards (2006), involves the ability to use a language in a variety of settings with a number of diverse individuals and for a range of distinct objectives.

Within the context of the KSA, Cullen (1998) clarified that, in order to transition from teacher-centred to learner-centred curricula, teachers across the country should be encouraged to adapt CLT in EFL classrooms. Nevertheless, Al-Seghayer (2005) argued that the CLT approach was not adopted within the EFL classroom as teachers showed no interest in using CLT with EFL learners. Reluctance can be present, even though evidence indicates that involving language learners in various communicative activities, as CLT, could improve communicative competence (Ellis & Ellis, 1994). Al-Seghayer (2005) and

Zaid (1993) suggested that perceiving reading and writing as the two vital language skills in EFL teaching and learning by teachers in Saudi Arabia could lead to resistance among teachers toward using CLT in classrooms. Alzaidi (2011) as well as Al-Seghayer (2015) noted that, among the reasons for the continued use of traditional teaching techniques, is a dearth of clear instructional resources on the implementation of CLT, insufficient training, high curricular loads, and learners' perceived lack of English proficiency; all of these factors could negatively affect teachers' implementation of CLT. Alzaidi (2011) and Al-Seghayer (2015) findings further indicated that Saudi EFL instructors had a strong theoretical foundation in the communicative language teaching approach and its concepts; even more surprisingly, instructors appeared to have a favourable perspective about the communicative language teaching approach. Nevertheless, in actual classroom practice, the CLT approach was not being adopted by the teachers. Instead, traditional teaching methods were still being used, just with new resources. Consequently, instead of assisting learners in using and practising the English language through active learning, teachers spent most of the class time concentrating on the delivering the content through the same old cognitive skills, namely those related to the ability to memorise acquired information.

Al-Seghayer (2015) argued that one consequence of the adherence to traditional teaching methods in the KSA is that context assessment is done only for summative purposes through oral and written testing. Al-Seghayer (2015) further argued that a dominating exam-oriented culture compels language instructors to do all possible to prepare pupils for final exams and, as a result, to focus exclusively on the topics and material covered by the tests. Indeed, Al-Sadan (2000) found that teachers' main concern was the number of their students who would pass exams. The adherence to traditional teaching methods is thus, at least in part, a consequence of the current practices and purposes of assessment in the Saudi context. Reconsidering current practices of English language assessment is therefore essential. Educational assessment can enhance and develop the educational processes and outcomes (Gordon & Rajagopalan, 2016). Accordingly, the following section addresses assessment in the Saudi education system, particularly EFL assessment in higher education.

2.6. English language assessment in higher education in the KSA

The primary objective of assessment in KSA is to measure the educational curriculum's outcome (Al-Saloom, 1987). Until recently, student assessment in universities has mainly

supported a model that is norm-referenced and summative (Darandari & Murphy, 2013). In Saudi Arabia, EFL assessment is viewed as a method of grading pupils according to their acquired knowledge and skills by the conclusion of course or programme of study, with formal examination serving as the primary means of gathering summative data (Darandari & Murphy, 2013). Heywood (2000) argued that traditional assessment techniques, as examinations, have been claimed to produce surface learning and raise student anxiety. This, according to (Hughes, 2003), is a result of the fact that examinations are mostly intended to assess students' performance on previously taught courses.

To illustrate, in higher education, student assessment is commonly undertaken in two parts: midterms and quizzes during the semester (with only 10–15% being dedicated to oral tests), which account for up to 60% of the assessment, and the final examination, which accounts for the other 40% of the assessment. The tests usually require students to recall memorised content and urge learners to be more concerned with their grades instead of the actual learning process; as a result, test-focused learning tends to target only the lower-order cognitive skills in Bloom's taxonomy (2001) and often puts too much emphasis on the final product (Simonson et al., 2000). These practices are largely inconsistent with an approach that would help learners meet the key objective of EFL teaching in Saudi Arabia; that is, to improve learners' intellectual, personal, and professional abilities. Additionally, Mohammed (2016) argued that tests were considered as an anxiety-inducing factor for Saudi EFL students.

Furthermore, Darandari and Murphy (2013) argued that traditional assessment approaches were not integrated into a learning experience to support students' learning; the emphasis on grading defines a curriculum that is centred on evaluation rather than on the processes and results of learning. Researchers have consistently suggested that the purpose of assessment should not be to only grade students' performance; rather, the purpose should be to offer invaluable information on the effectiveness and suitability of the learning and teaching process for students and teachers, as well as the quality of curriculum creation and implementation (see, for example, Berry (2008); Lambert and Lines (2013); Natriello (1987); Newton (2007, 2010); Sadler (2010). Chapter 3 addresses these issues in more depth. Darandari and Murphy (2013) assessment, when incorporated into a plan for improving the quality of the teaching and learning process, should be created simultaneously with educational practice, curriculum development and design. Nonetheless, there is no evidence of this strategy in the KSA at the system level, university level, or individual classroom level (Darandari & Murphy, 2013).

Despite efforts in the KSA to encourage adoption of student-centred approaches, traditional methods for student assessment are still dominant in HE (Al-Seghayer, 2022; Almossa, 2021). The loyalty to traditional assessment is also an issue in the UK HEIs, as observed by Hargreaves (2006); also in the UK, there is an emphasis on traditional assessment methods, specifically examinations, as the mainstay of all assessment despite the rich and varied approaches to assessment that are now available. In 2005, the National Commission for Academic Accreditation and Assessment (NCAAA) was formed by Saudi Arabia, demonstrating the country's strong commitment to improve, as highlighted by Almusallam (2007). The new commission has promoted a new paradigm for higher education, which concentrates on learning outcomes and stresses the importance of planning and designing assessment around those learning outcomes (Al-Musallam, 2007; Darandari & Murphy, 2013). Consequently, attitudes have started to shift, and new strategies have been created to offer more comprehensive and flexible assessment approaches capable of achieving a broader variety of objectives (Darandari & Murphy, 2013). It is noteworthy that, over the years, NCAAA was devoted to teaching with a variety of published documents and workshops supported by quality assurance deans in Saudi institutions (NCAAA, 2015). With the emphasis on teaching, these materials and workshops appeared to disregard assessment (Almossa, 2021).

Though traditional methods for student assessment remain dominant, in the last few years, teacher-centred assessment approaches have been challenged by many Saudi academics in HEIs, with some Saudi lecturers and researchers considering alternative assessment approaches such as portfolios and self and peer assessment (see, for example, Alotaibi (2019); Alahmadi et al. (2019); Al-Abdullatif (2020).

In an attempt to comprehend the role and impact of self-assessment on EFL learners' self-regulated learning and critical thinking, the literature review that follows in Chapter 3 explores the literature on self-assessment, self-regulated learning, and critical thinking.

Chapter 3 : Literature Review

3.1. Introduction

Chapter 2 sets the context for this study in the education system in KSA. The chapter explored the historical journey of English language teaching and its impact on current English language teaching and assessment in HE in the Kingdom of Saudi Arabia. Chapter 3 now moves beyond the specific context of Saudi Arabia and reviews the literature on assessment, self-regulated learning, and critical thinking more broadly.

Over time, consistent themes have emerged in the EFL literature as to what makes learning successful. Holec (1981), in a study carried out in the field of foreign language learning, was the first to introduce the centrality of autonomy and self-regulated learning; in other words, the ability of an EFL learner to manage and master learning within tasks is crucial to high-quality learning. Building on Holec's work, Roghanizadeh (2011), Ghanizadeh and Mirzaee (2012), and Kamgar and Jadidi (2016) presented empirical evidence to suggest that effective learning in EFL requires not only self-regulated learning but also critical thinking. In accordance with Gardner (2012), the present research project is concerned with the influence of assessment practices that have the potential to have a substantial impact on future learning. However, before exploring how self-regulated learning and critical thinking may be enhanced through self-assessment, it is necessary to appraise self-regulated learning and critical thinking as concepts.

For the purpose of understanding the role and impact of self-assessment on EFL learners' self-regulated learning and critical thinking, the literature review is divided into two sections. The first section introduces and discusses the concepts of self-regulated learning and critical thinking and considers the link between the two. The second section aims first to provide and discuss the background for the term *self-assessment* by exploring two different sets of assessment lenses: (1) *summative assessment* and *formative assessment* and (2) *Assessment of Learning (AoL)*, *Assessment for Learning (AfL),* and *Assessment as Learning (AaL)*. The chapter next investigates the concept of self-assessment, including its definitions and its contextualisation in different educational paradigms, in particular the constructivist and the social constructivist paradigms. The link between conceptualisations and practices of self-assessment used in EFL will then be considered. This chapter then concludes by exploring the potential of self-assessment as a means to enhance critical thinking and self-regulated learning in an EFL context.

3.2. Self-regulated learning and critical thinking in education

The relationship between self-regulated learning and critical thinking is significant to this research and is discussed in depth in Section 3.2.3. Before considering the relationship between the two concepts, this section outlines the origins and literature surrounding self-regulated learning and critical thinking. The following three sections aim to offer an evidence-based overview of self-regulated learning and critical thinking, including their significance in education and the relationship between the two concepts.

3.2.1. Self-regulated learning

The topic of self-regulated learning, as well as self-regulation in general, emerged from the area of autonomous learning. The concept of learner autonomy gained substantial attention in education around the 1970s (Little, 1991), with the primary concentration on the abilities of learners to take responsibility for their own learning and lifelong learning (Griffiths & Parr, 2001). The term *autonomous learner* is difficult to define. After reviewing titles of books on autonomy, Pemberton et al. (1996) argued that "different terms are often used to refer to the same thing and the same term is often used to mean different things" (p. 2). This research follows Holec's (1981) view of autonomy as his work, *Autonomy in Foreign Language Learning (1981),* was a fundamental basis for other researchers in the context of foreign language learning. Holec (1981) began by defining learner autonomy as the ability to be responsible for personal learning, noting that this ability must be learnt, it is not an inborn ability, either through natural means or via systematic and deliberate formal learning. Holec (1981) further explained that being accountable for every choice made in relation to all aspects of learning is what it means to take responsibility for personal learning. Holec's work sought to promote learner autonomy through developing in learners the abilities that would allow learners to become more responsible while contributing to the society in which they live. Learner autonomy thus increases learners' potential to play an active role in and take control of every stage of their learning, from goal setting to evaluation. It is well documented in the work of Sierens et al. (2009), Andrade and Bunker (2009) and Murray (2014) that gaining autonomy, which entails the added responsibility of taking charge of one's own learning, requires the application of self-regulatory processes. Rowe and Rafferty (2013) identified numerous self-regulatory processes, including recognising knowledge gaps, making informed decisions about personal educational requirements, managing time effectively in terms of how and when to study, planning, engaging in reflective tasks, and maintaining sustained

motivation to accomplish tasks successfully. Zimmerman and Schunck (2001) characterised self-regulated learning as the act of controlling and mastering personal learning within tasks.

However, Dinsmore et al. (2008) and Oppong et al. (2019) pointed out that self-regulation is not the same thing as self-regulated learning. This distinction exists despite the overlap in the two terms themselves: *self-regulation* and *self-regulated learning*. Both terms are often used interchangeably in educational literature, but the distinction matters, as understanding the distinction can assist educators in identifying the scientific and theoretical basis for their practices.

3.2.1.1. Self-regulation (SR) versus self-regulated learning (SRL)

Identifying the distinction between self-regulation and self-regulated learning is necessary to the basis for this study. Initially, self-regulation was significantly influenced by the work of Albert Bandura, whose initial publications were issued in the 1970s. In the work of Bandura (1977), self-regulation was not viewed as having an educational or cognitive enterprise; it was instead the process by which a person controls one's own behaviour, actions, thoughts, or emotions in the context of life in general, not within the learning context. It concerns issues such as addictions and anger management. It would only later be suggested by Carver and Scheier (1981/2012) that self-regulation can be related to learners through its idea of goal-setting and feedback, since people's behaviours are guided by goals and feedback. This suggestion put an emphasis on the use of self-regulation in a learning context (Carver & Scheier, 2017). Self-regulation has continued to develop and emerge in research such as in the work of Zimmerman and Bandura (1994) and Zimmerman and Schunck (2001). Developments in research have linked self-regulation to motivation, specifically self-efficacy, namely one's belief about their personal ability to succeed in a task or in any circumstance (Zimmerman & Schunck, 2001; Zimmerman & Bandura, 1994). As the academic interest in self-regulation increased, Zimmerman and Bandura (1994) defined self-regulation in relation to learning; their more recent definition, which refers to self-regulation of learning, addresses how learners' self-generated thoughts, feelings, and behaviours can be organised systematically towards achievement of educational goals. Zimmerman and Bandura's (1994) definition of self-regulation in relation to learning reflects Holec's (1981) view of learner autonomy and the idea of learners taking control of their own learning in a systematic way. Thus, the increasing emphasis on self-regulation in academic contexts seems to have had a significant role in

the development of a new term: *self-regulated learning.*

Self-regulation has been seen as an umbrella term under which the term *self-regulated learning* falls. The term self-regulated learning was introduced in the 1980s and increased in prominence in the 1990s (Lajoie, 2008). According to Schunk (2001), self-regulated learning is characterised by learners' self-generated thoughts and behaviours being systematically concentrated on achieving study objectives. Self-regulated learning has mostly been conceptualised to contain regulation within an academic field. In contrast, self-regulation is not necessarily about learning at all, and it is only sometimes applied in academic contexts or cognitive tasks. Zimmerman (2002) argued that self-regulated learners can be perceived as being guided by personally and individually established goals and task-related strategies; self-regulated learners are, therefore, conscious of their strengths and weaknesses, and proactive in their learning. He argued that through monitoring their behaviour in relation to their goals and reflecting on the effectiveness of their performance, self-regulated learners increase their sense of self-esteem and motivation, which eventually results in improved performance and learning techniques. Zimmerman (2002) also argued that self-regulated learners have a greater possibility of attaining academic achievement and have an optimistic perspective of the future resulting from their motivation and adaptable learning techniques.

Self-regulated learning is primarily considered a phased, intentional, complex, and goal-directed process containing cognitive, metacognitive, motivational, affective, emotional, social, and contextual components that are always tied to learning (Muis, 2007; Oppong et al., 2019). Muis et al. (2018) noted that self-regulated learning can reveal the active and continues progression of learning and learners' knowledge. Similarly, Winne and Hadwin (1998) and Oppong et al. (2019) stated that self-regulated learning could extend students' experience by taking students' assumptions and perceptions of their competence to complete a task using the available resources into consideration. Self-regulated learning can also be argued to incorporate learners' feelings about tasks' engagement as well, which is referred to as personal epistemology (Oppong et al., 2019; Winne & Hadwin, 1998).

By taking personal epistemology into consideration, self-regulated learning adds a new and important concept to self-regulation (Oppong et al., 2019). Generally, personal epistemology has been centred on an individual's epistemic views; that is, how people think about knowledge and what they believe about it, as those factors could affect cognitive processing, and conceptual change learning as well as affect strategy

implementation and understanding (Hofer, 2008). In other words, personal epistemology is the act of reflecting on or being critical of one's own thought procedures or methods (Mason et al., 2010). Self-regulated learning includes four aspects of personal epistemology: epistemic cognition, epistemic metacognition, epistemic motivation, and epistemic emotions (Mason et al., 2010). Epistemic cognition includes the learners' knowledge and ideas about their individual learning and problem-solving processes, as well as about learning and problem-solving in general, typically in a specific context (Mason et al., 2010). Epistemic metacognition addresses how learners' general planning, monitoring, evaluating, and revising of learning and problem-solving methods, in addition to the sequencing of these strategies, are influenced by their knowledge and beliefs (Mason et al., 2010). Epistemic motivation engages learners in critical and deep thinking (Scholten et al., 2007) and enables learners to develop and maintain well-informed conclusions independently (Amit & Sagiv, 2013). Epistemic emotions include the feelings that emerge when learners' primary attention is on knowledge and knowing (Boekaerts & Pekrun, 2015). Therefore, due to the concentration on personal epistemology, self-regulated learning tends to emphasise self-motivation, self-awareness, and behavioural skills as essential components of successful knowledge use; that is, self-regulated learning exceeds the need for skill-specific knowledge. Thus, personal epistemology is the major difference between self-regulation and self-regulated learning.

For the purpose of this research, a decision between these two concepts, self-regulation and self-regulated learning, had to be taken in order to form the research basis. As indicated above, self-regulation is not tied to the learning context in particular; it is more concerned with the context of life in general. Self-regulated learning, in contrast, is constantly connected to learning and constrained by the environment or the situation of learning (Bandura, 1977; Oppong et al., 2019; Winne & Hadwin, 1998). Moreover, emphasis should be placed on learners' assumptions and ideas about their competence for a particular task, as well as learners' emotions about their involvement with the activities, as all influence their experiences and development; as such, these elements relate to self-regulated learning, not to self-regulation, since these elements are rooted in students' personal epistemology, as noted by Winne and Hadwin (1998).

One of the main purposes of the present research is to explore and understand EFL learners' perceptions and experiences of the impact of self-assessment. Accordingly, the theory of self-regulation is not suitable and does not fit the purpose of this research. Self-regulated learning, however, has been recognised as an effective technique for overcoming

underachievement among students who possess the capacity to succeed at very high levels. Indeed, underachievement may be ascribed to a variety of factors, including a lack of early definition of the learners' task, insufficient motivation or efficacy, or unpleasant feelings associated with the environment or task (Oppong et al., 2019). Thus, self-regulated learning is more suitable for the purpose of this research (see Table 3-1 for a summary of self-regulation and self-regulated learning key features).

Table 3-1 Self-regulation versus self-regulated learning

	Self-regulation (SR)	**Self-regulated learning (SRL)**
Definition	The process by which a person controls one's own behaviour, actions, thoughts, or emotions in the context of life in general, not within the learning context (Bandura, 1977).	self-regulated learning is characterised by learners' self-generated thoughts and behaviours being systematically concentrated on achieving study objectives (Schunk, 2001). It has mostly been conceptualised to contain regulation within academic field (Schunk, 2001).
Key features	It can be related to learners through its idea of goal-setting and feedback (Carver & Scheier, 2017). Developments have linked self-regulation to motivation, specifically self-efficacy (Bandura, 1988; Zimmerman & Schunck, 2001; Zimmerman & Bandura, 1994)	It is goal-directed process containing cognitive, metacognitive, motivational, affective, emotional, social, and contextual components (Muis, 2007; Oppong et al., 2019). Self-regulated learning adds an important concept to self-regulation that is personal epistemology (Oppong et al., 2019). Consequently, self-regulated learning tends to emphasis self-motivation, self-awareness, and behavioural skills as essential components of successful knowledge use; that is, self-regulated learning exceeds the need for skill-specific knowledge.

3.2.1.2. Self-regulated learning consensus: Phases, components, and processes

Research in the area of self-regulated learning is broad, yielding a number of theories and models seeking to characterise and differentiate effective learner characteristics (see, for example, Bandura (1991); Boekaerts and Niemivirta (2000); Pintrich (2000); Zimmerman (1990)). Nevertheless, there is a widespread agreement on the phases and component of SRL (Bandura, 1991; Beishuizen & Steffens, 2011; Manso-Vázquez et al., 2014, 2018; Pintrich, 2000; Zimmerman & Schunck, 2001). Manso-Vázquez et al. (2018) proposed an integrated model that incorporates the three most often used models in self-regulated learning: Zimmerman (2002) model, Pintrich (2004) general model, and Winne and Hadwin (1998) information processing model. The integrated model of Manso-Vázquez et al. (2018) consists of three distinct phases and five distinct areas. The three phases are (1) forethought, planning, and activation; (2) performance, monitoring, and control; and (3) evaluation, reflection, and reaction. Cognition, metacognition, motivation, behaviour, and social and environmental context are the five areas in the integrated model (see Figure 3-1 for details of the phases, areas, and learning strategies). The work of Pintrich (2000), Bol and Garner (2011), Schunk and Zimmerman (2012), and Panadero (2017) has affirmed that SRL incorporates several crucial cognitive, metacognitive, motivational, and behavioural affective aspects that influence learning. Self-regulated learning is therefore considered to be a very important area of research (Panadero, 2017). The areas in self-regulated learning are discussed in Section 3.3.5 in relation to self-assessment. Furthermore, the implementation of the integrated model of Manso-Vázquez et al. (2018) is presented in Chapter 4.

Figure 3-1 Integrated self-regulated learning model (adapted Manso-Vázquez et al., 2018, p. 42469)

3.2.1.3. Measuring self-regulated learning skills

The measurement of self-regulated learning has been a significant research issue in the area of SRL. The different proposed types of measurement include mainly self-reporting and event measurement. Zimmerman (2008) emphasised the self-reporting technique, or the use of questionnaires, surveys, and interviews. This technique mainly focuses on the perceptions and ideas of the learners, i.e. on learners' perspectives of self-regulated learning. In contrast, Winne and Perry (2000) proposed event measures such as traces and thinking aloud protocols for measuring the process of self-regulated learning. Panadero et al. (2012) and Winne and Jamieson-Noel (2002) found that self-reporting as a technique for measuring self-regulated learning may not always be accurate, since students may not always report their true usage of methods. Additionally, Boekaerts (1997) argued that self-reporting does not necessarily record changes in student use of strategies. Therefore,

Boekaerts and Corno (2005) recommended the triangulation of self-regulated learning data; that is, not relying on one measure. On the same note, Panadero et al. (2012) emphasised the need to combine self-reporting measures with event measures. The techniques used in measuring self-regulated learning are discussed in Chapter 4 in relation to this research.

With the concept of self-regulated learning clarified, the next section provides an evidence-based overview of the concept of critical thinking.

3.2.2. Critical thinking

Adopting critical thinking as an educational goal could promote students' autonomy and prepare students to succeed in life. The term *critical thinking* refers to the concept of reflective thinking, which originated as a way to describe educational objectives by the American philosopher John Dewey (1910/1997). Dewey (1910/1997) defines critical thinking as "active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends" (p. 9). However, the definition of critical thinking is contested and has changed over time. A number of definitions have been proposed. Ennis (2016) lists 17 definitions of critical thinking, 13 of which are philosophically oriented scholarly definitions. Critical thinking in educational contexts is generally given as a "programmatic definition" (Scheffler, 1960, p. 19). Building on Scheffler's (1960) research, Ennis (1969) argued that a programmatic concept or definition takes the form of a definition, but it is not only a definition. Ennis (1969) defined a programmatic concept as a proposition or suggestion to adopt and implement a programme or a particular perspective. The definition of critical thinking should, thus, perhaps also be related to a programme of what ought to be done to encourage critical thinking in practice. Similarly, Hitchcock (2018b) argued that, because the attainment of an educational goal through a practical programme characterises critical thinking as part of education, it is far more beneficial to identify criteria and standards of critical thinking than to provide a definition in a single sentence. As Hitchcock (2018b) further explained, these criteria and standards include the knowledge, abilities, and disposition of critical thinking. Hitchcock (2018b) suggested that recognising, adopting, and implementing these criteria and standards should be considered as the actual educational goal.

Despite the variation in definitions of the concept, there is a widespread agreement that

critical thinking contains cognitive abilities and dispositions. Ennis (1987), Facione (1990), and Halpern (1998, 1999) have argued that critical thinking is more than the successful use of a skill; it is also an attitude or disposition to recognise when a skill is needed as well as the commitment and willingness to invest the mental effort required to apply this ability. Paul and Elder (2006), Bailin and Battersby (2016), and Hamby (2014) proposed using the *virtues* to refer to the critical thinking dispositions. Virtues, in context of critical thinking, are not moral virtues but intellectual virtues, as presented and discussed by Zagzebski and Zagaebski (1996) and Turri et al. (2017). These virtues concern a person's modes of thought rather than ways of interpersonal interactions (Turri et al., 2017). On this basis, it can be argued that critical thinkers have abilities and dispositions that, when appropriate, can encourage and lead individuals to think critically. Hence, the following sections explore both the cognitive ability and dispositions or intellectual virtues of critical thinking.

3.2.2.1. Components of critical thinking

3.2.2.1.1. Cognitive ability

As Ku and Ho (2010a) clarified, an individual's cognitive ability is characterised by the ability to not only understand a problem but also engage in rational decisions based on the adoption of cognitive skills. Cognitive abilities are central to critical thinking (Ku & Ho, 2010a). Hitchcock (2018b) identified these cognitive abilities as observational abilities, which involve making careful and accurate observation; emotional abilities, which involve identifying one's own emotional commitments and responses, as well as those of others; questioning abilities, which involve formulating inquiries; imaginative abilities, which involve generating possible explanations; inferential abilities, which involve establishing conclusions from provided data and determining the degree of confidence with which one's own or others' conclusions follow; experimenting abilities, which involve designing and executing experiments; consulting abilities, which involve finding and appraising information; argument analysis abilities, which involve identifying and evaluating arguments; and judging and deciding skills, or the capacity to recognise the conclusion or judgement that the available facts and reasoning support. Hitchcock (2018b) argued that identifying and focusing on cognitive abilities is a beneficial first step before establishing educational objectives. Setting educational objectives, in turn, is a beneficial first step before developing strategies to assist learners in achieving the objectives and, naturally, prior to developing methods for assessing the degree to which learners have reached goals

(Hitchcock, 2018b). In short, recognising and targeting certain cognitive abilities is key in developing the process and strategies used for developing learners' critical thinking. Nevertheless, as Section 3.2.2 explored in detail, those skills and attitudes or dispositions are significant aspects of good critical thinking performance. The following section explores dispositions, or the intellectual virtues of critical thinking.

3.2.2.1.2. Dispositions

There is considerable agreement amongst researchers that, along with abilities, critical thinking also entails dispositions (see, for example, Ennis (1987); Facione (1990); Bailin and Battersby (2016); Halpern (1998); Hamby (2014); Paul and Elder (2006)). The term *dispositions* is used broadly to refer to the habits of mind and attitudes contributing causally to being able to think critically and influence the patterns of a person's intellectual activity (Hitchcock, 2018b). Critical thinking dispositions are identified by Ku and Ho (2010a) as enjoyment of thinking, an open attitude, a careful approach in thinking, and a mindset for truth; these dispositions are essential for a person to reach sound judgements. As Facione (1990) explained, there are two forms of dispositions that may be helpfully distinguished: those that contribute causally to the initiation of critical thinking on an issue, i.e. initiating dispositions, and those that contribute causally to an individual's ability to think critically once one has started, i.e. internal dispositions.

Hitchcock (2018b) adopted the approach of appraising variables that could hinder the critical thinking capabilities of individuals. By doing so, he established that the initiating dispositions include pursuing truth, interest in finding evidence to support one's own perspectives, confidence in reason, readiness to delay judgement, open-mindedness, tenacity, self-assurance, inclination to investigate, and attentiveness. Pursuing truth and interest in finding evidence to support one's own perspectives promotes critical thinking by encouraging the learner to move past their initial subjective perspective on a subject. Confidence in reason, which leads to respect for the procedure of rational investigation, also characterises the development of critical thinking, discouraging suspicion of reason. Readiness to delay judgement when considering various explanations and options is indicative of an inclination towards critical thinking, which can be hindered by swift decision-making. Open-mindedness counters inflexibility, which is a hindrance to critical thinking. Tenacity, especially in intellectual pursuits, makes an inclination towards critical thinking, and self-assurance counters any stifling of independent thought that may arise due to limited confidence in one's own critical thinking capacities. The inclination to

investigate stems from an inner motivation to pursue meaningful inquiry into something. Finally, attentiveness helps to ensure an individual does not overlook how critical thinking about something is required.

Moreover, as Hitchcock (2018b) clarified, certain initiating dispositions are also internal dispositions, for example readiness to delay judgement and open-mindedness. Nevertheless, Hitchcock explained that internal dispositions are associated with cognitive approaches or trends that causally influence effective critical thinking after it has been initiated. Furthermore, Hitchcock (2018c) related that internal dispositions, like an inclination to endure in completing a difficult activity, may be impetuses to endeavour with, or amend, the critical thinking procedure. Alternatively, tenets of effective thought, such as establishing an issue with clarity and staying focused on that issue, provide the basis of numerous other internal dispositions for critical thinking.

There is considerable consistency amongst different researchers who have established comparable sets of dispositions pertinent to critical thinking abilities (Bailin et al., 1999; Ennis, 1987; Facione, 1990; Halpern, 1998; Paul & Elder, 2006). The following are commonly referred to as critical thinking dispositions: an inclination to acknowledge and respect other people's perspectives (Bailin et al., 1999); flexibility (Facione, 1990; Halpern, 1998); a desire to possess knowledge (Ennis, 1987; Facione, 1990); curiosity (Bailin et al., 1999; Facione, 1990); an inclination to pursue reason (Bailin et al., 1999; Ennis, 1987; Paul & Elder, 2006); fair-mindedness (Bailin et al., 1999; Facione, 1990); and open-mindedness (Bailin et al., 1999; Ennis, 1987; Facione, 1990; Halpern, 1998).

Nonetheless, critical thinking is considered to be a combination of both dispositions and abilities, as Section 3.2.2 explained. Ku and Ho (2010b) agreed with this position, as they argued that cognitive abilities and dispositional components, in addition to metacognitive components, which will be discussed in the following section, are all required for effective critical thinking. Consequently, both dispositions and capabilities for critical thinking must be fostered. Additionally, there is general agreement that while direct instruction may facilitate the development of abilities, dispositions are better referred to as attitudes, thus long-term engagement in learning contexts receptive to reflection and debate is necessary for the dispositions' development (Ennis, 1987; Facione, 1990; Halpern, 1998, 1999).

3.2.2.1.3. Measuring components of critical thinking

Critical thinking standardised tests have been created to determine an individual's degree of possession of certain dispositions and skills (Hitchcock, 2018b). Experimental educational interventions have demonstrated that education can promote critical thinking abilities and dispositions, as measured by standardised tests (see, for example, Bixler et al. (2015); Naber and Wyatt (2014); Nelson et al. (2018); Webster and Willett (2019) Maryam et al. (2021)). While the abilities can be identified directly, the dispositions need to be identified indirectly, since dispositions relate to habits of mind and attitudes; indirect identification may involve considering what factors contribute to, or impede exercise of, the dispositions (Hitchcock, 2018b). Consequently, critical thinking skills are more easily quantifiable than critical thinking dispositions (Hitchcock, 2018a). The following eight currently available standardised tests have been designed to measure critical thinking skills as highlighted by (Hitchcock, 2018b) including: the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1994); the Cornell Critical Thinking Tests Level X and Level Z (Ennis, Millman, & Tomko, 1985, 2005); the Ennis-Weir Critical Thinking Essay Test (Ennis & Weir, 1985); the California Critical Thinking Skills Test (Facione, 1992); the Halpern Critical Thinking Assessment (Halpern, 1998); and the Collegiate Learning Assessment (Council for Aid to Education, 2017). In the field of language learning, the California Critical Thinking Skills Test (Facione, 1992) and the Watson-Glaser Critical Thinking Appraisal (Watson & Glaser, 1994) are the two most commonly used tools for measuring critical thinking skills in the field of language learning. In the current research, the Watson-Glaser Critical Thinking Appraisal test was used to measure the students' critical thinking abilities. This test is discussed in more detail in Chapter 4. It is more difficult to measure dispositions with the use of a multiple-choice format – as noted, dispositions concern habits of mind – yet, some standardised tests do exist for that purpose (Hitchcock, 2018a) such as the California Critical Thinking Dispositions Inventory. Nevertheless, a more effective method of assessing critical thinking dispositions would be to explore how individuals behave when confronted with situations that reveal their dispositions (Hitchcock, 2018a). Therefore, in the current study, the student participants were involved in a self-assessment intervention followed by interviews, which could reveal their dispositions and indicate if the students exhibited a good critical thinking performance. According to Ku and Ho (2010b), a good critical thinking performance involves cognitive ability, including both dispositional and metacognitive components.

Regarding metacognitive components, limited research has investigated the link between metacognitive components and critical thinking, even though this link has been discussed as an important factor affecting critical thinking (Ku & Ho, 2010b). The main aspect of the link is that metacognitive components are believed to elicit behaviours in pupils that allow them to monitor and regulate their thinking processes and abilities, and dispositions can arise in the process (Ku & Ho, 2010b). In other words, a critical thinker oversees their thinking processes thanks to the use of metacognitive strategies, which enable the thinker to exert control.

The most commonly recommended metacognitive approaches for critical thinking fall into three groups: planning, monitoring, and evaluating (Ku & Ho, 2010b). Planning strategies include those aimed at establishing processes that guide thinking, and include the selection of suitable tactics and resource allocation (King, 1991; Schraw, 1998, cited in Ku & Ho, 2010b). Monitoring is a term that refers to a continuous awareness of task understanding (Schraw, 1998, cited in Ku & Ho, 2010b). Monitoring strategies thus include verifying task information to ensure understanding, directing attention to critical concepts, and highlighting informational ambiguities (Luckey, 2003; Swartz, 2003, cited in Ku & Ho, 2010b). Evaluating strategies include the evaluation and correction of an individual's cognitive processes (Facione, 1990, cited in Ku & Ho, 2010b). They involve evaluating one's own logic, objectives, and conclusions (Schraw, 1998, cited in Ku & Ho, 2010b) as well as making revisions when necessary (see Table 3-2 for a summary of critical thinking components). The relationship between metacognitive strategies and critical thinking implies that self-regulated learning and critical thinking are linked, a relationship that is discussed in the following section.

Table 3-2 A summary of critical thinking components

Components of critical thinking	Cognitive ability	Dispositions/ Virtues		Metacognitive
Definition	Cognitive ability is characterised by the ability to not only understand a problem but also engage in rational decisions based on the adoption of cognitive skills.	Refer to the habits of mind and attitudes that contribute causally to being a critical thinker and influence the patterns of a person's intellectual activity (Hitchcock, 2018a).		Metacognitive components are believed to elicit behaviours in pupils that allow them to monitor and regulate their thinking processes and abilities, and dispositions can arise in the process (Ku & Ho, 2010b)
Components	Hitchcock (2018b) identified the following cognitive abilities: • observational abilities • emotional abilities • questioning abilities • imaginative abilities • inferential abilities • experimenting abilities • consulting abilities • argument analysis abilities	**<u>Initiating dispositions</u>** are those that contribute causally to the initiation of critical thinking on an issue (Facione, 1990). Hitchcock (2018b) identified the following initiating dispositions: • pursuing truth • confidence in reason • readiness to delay judgement • open-mindedness • tenacity • self-assurance • attentiveness • Hitchcock (2018b)	**<u>Internal dispositions</u>** are those that contribute causally to an individual's ability to think critically once one has started (Facione, 1990). • Flexibility (Facione, 1990; Halpern 1998). • Curiosity (Bailin et al., 1990; Facione, 1990). • Pursuing reason (Ennis, 1985; Paul, 1992; Bailin et al., 1999). • Desire to possess knowledge (Ennis, 1984; Facione, 1990). • fair-mindedness (Facione, 1990; Bailin et al., 1999) • open-mindedness (Ennis, 1985; Facione, 1990, 2000; Halpern, 1998; Bailin et al., 1999). • Willingness to acknowledge and respect other people's perspectives (Bailin et al., 1999)	Metacognitive techniques for critical thinking fall into three groups (Ku & Ho, 2010b): • Planning strategies • Monitoring strategies • Evaluating strategies

3.2.3. The relationship between self-regulated learning and critical thinking

Several researchers have suggested that critical thinking and self-regulated learning are interrelated, so improving one leads to improving the other; both have a positive effect on learners' learning (Butler, 2002; Dickinson, 1987; Dunn et al., 2014; Ghanizadeh & Mirzaee, 2012; Kahrizi & Farahian, 2014; Phan, 2010; Uyar et al., 2018; Watson, 2002; Zimmerman, 2002). Dickinson (1987) and Zimmerman (2002) argued that self-regulated learning and critical thinking are interrelated, because a self-regulated person who is critical in their thinking is capable of learning more efficiently and quickly than an individual who lacks these qualities. Phan (2010) discussed that self-regulated learning is facilitated by critical thinking as a cognitive skill. Critical thinking requires the use of higher-order learning techniques, one of which is self-regulation, at least when self-regulation is employed to analyse data and assess classroom activities (Phan, 2010). In the same vein, Butler (2002) argued that improving self-regulated learning is linked to the development of students' capacity to analyse tasks, and that to analyse tasks students need to employ critical thinking. Likewise, Ghanizadeh (2011) indicated that critical thinking skills contribute to the long-term improvement of self-regulation. Accordingly, it appears that, through active involvement in the learning process, learners develop self-regulated learning and critical thinking. In turn, self-regulated learning together with critical thinking leads learners to acquire a sounder understanding of their learning objectives, to have greater recognition of their own knowledge and abilities, and to embrace the development of further learning strategies. Engaging students in practices as a means of enhancing self-regulatory skills and critical thinking is, therefore, significant.

Evidence suggests that self-assessment, as an example of a practice, is significantly positively correlated with self-regulated learning and critical thinking. Paul et al. (1997) argued that self-assessment is a key aspect for critical thinking, and that critical thinkers are those who have mastered self-assessment. They contended that a critical thinker is capable of monitoring, analysing, judging, and selecting the most effective way of thinking; thus, a good critical thinker is, in effect, a good self-assessor (Paul et al., 1997). Similarly, Smith (1997) emphasised the significance of self-assessment, stating that self-assessment enables learners to develop into independent learners, and their capacity for critical thinking is defining characteristic of independent learners. The argument here is that the idea of self-assessment fosters critical thinking. However, it is important to note that Smith's (1997) work linked autonomous learning with critical thinking. In the area of

autonomous learning, Campbell et al. (1998) stated that self-assessment instils in learners a stronger feeling of ownership over their work, a more passionate attitude toward learning, and an improved capacity for higher-order thinking; self-assessment thus, they concluded, increases students' self-regulation in learning. All the aforementioned evidence suggests that self-assessment can promote both self-regulated learning and critical thinking, as self-assessment is generally concerned with giving opportunities for students to self-assess their achievement and concentrate on their learning process. This thesis uses the term *self-assessment* throughout. Thus, this thesis intends to investigate the term's origin and the theoretical underpinning of self-assessment before examining the influence of self-assessment on EFL learners' self-regulated learning and critical thinking.

3.3. Self-assessment

Assessment is of paramount importance in measuring, supporting, and enhancing learning; tracking progress; and informing the teaching process for certification, selection, and accountability purposes (Gardner, 2012). The focus of this research is the importance of assessment in enhancing learning. Helping and supporting learners to improve their learning is one of the fundamental ideas of assessment; thus, assessment is essential to promoting learning and eventually attainment (Brink, 2017). Self-assessment, in particular, is a powerful strategy for involving learners in their own learning and increasing their awareness in everyday educational practice.

This thesis focuses on the effect of self-assessment on self-regulated learning and critical thinking within an EFL context, namely, the Saudi context, but the literature search was not restricted to EFL contexts alone. It also included ESL contexts and English-speaking countries in order to find any evidence that would be related to this thesis. The major online research databases ERIC, JSTOR, and ProQuest, were used to obtain relevant studies using the keywords: Self-assessment OR Self-evaluation AND critical thinking OR self-regulation. The additional search terms "speaking", "ESL", and "EFL" were included to limit the search. Literature continued to be searched throughout the writing of this thesis.

The following sections overview the concept of self-assessment. First, the concept of assessment is introduced, and then the relationship between self-assessment, self-regulated learning, and critical thinking is discussed.

3.3.1. The concept of assessment

Assessment is central to education, as it is one of the three major message systems: curriculum, pedagogy, and assessment. However, *assessment* is also a contentious term; there is a lack of agreement on its definition (Taras, 2005). The word *assess* originates from 'to set by' or 'to sit down beside' (e.g. as an assessor or assistant/judge) from Old French and Late Latin languages (Oxford English Dictionary, n.d., cited in Wyse et al. (2015, p. 3). The central idea in this definition is the use of assessment to support learning, yet often the term *assessment* is associated more directly with the idea of judgement. This confusion between the learning and judgement functions of assessment can be illustrated in how the terms *assessment* and *evaluation* are sometimes used interchangeably in the field of education (Taras, 2005). This inconsistency emphasises the complexity of defining assessment and, thus, arguably justifies the shift in focus to forms of assessment instead of simplistic definition. Wiliam (2011) highlighted the increasing tendency to view the processes and activities intended to guide learners towards goals as approaches of assessment.

3.3.1.1. Developing conceptualisation of assessment

Initially, assessment was conceptualised as *summative assessment* and *formative assessment.* The terms summative assessment and formative assessment were introduced by Bloom (1971) as two kinds of assessment in students' learning. The terms were originally generated from the terms *summative evaluation* and *formative evaluation* by Scriven (1967). Scriven (1967) introduced the terms, summative evaluation and formative evaluation, pertaining to the evaluation of educational programmes comprising a curriculum, instructional materials, and general teaching strategies. Scriven (1967) used summative evaluation when referring to the final evaluation of a curriculum or programme; formative evaluation, in comparison, referred to evaluation throughout the development and implementation of new curricula with the aim of executing revisions and improvements before the final evaluation.

In the work of Bloom et al. (1971), summative assessment is seen as "judging, grading and certifying what the learner had achieved at the end of a course or programme that is usually in the form of examinations" (p. 20). In agreement with Bloom et al.'s perspective of summative assessment, Andrade and Cizek (2010) indicate that summative assessment involves the administered examinations in particular, predetermined, times throughout the

academic year, with the main purposes of evaluating the teaching process's success and classifying pupils according to the degree to which they have shown competence in a certain topic. This view of summative assessment seems to assume students' disengagement in the process of assessment. In comparison, formative assessment seems to stress the significance of student engagement in the learning process, as highlighted in the work of Bloom et al. (1971).

Bloom et al. (1971) described formative assessment as assisting the process of teaching and learning while remaining flexible and open to amendment. In other words, Bloom et al. argued that formative assessment aims to provide instructors with the possibility to intervene "during the formation of the student" (p. 20). Similarly, Ramaprasad (1983) stated that formative assessment is basically feedback to both the instructor and students on their current level of knowledge and skill development in order to identify the best course of action (Harlen & James, 1997). Mats Björkman's (1972) notion of "feedforward" seems relevant here, as it emphasises the importance of information exchange flow between the instructor and the student, which can take place through feedback to inform future efforts (Sadler, 2010). This notion involves providing information to be used to alter and develop the teaching and learning processes. Notably, though Bloom et al. (1971) claimed both kinds of assessment were to be used together, working in harmony, they described formative assessment as being used to improve summative assessment results; that distinction could indicate that formative assessment is less important than summative assessment.

In the early 2000s, Black and Wiliam (2003) argued that summative assessment was becoming prominent due to the external pressure for certification and accountability. Harlen and Deakin Crick (2003) showed that setting objectives based on test results could lead to a variety of practices that limited students learning experience and negatively affected students' learning motivation, self-esteem, effort, and achievement. Harlen (2005); Harlen and Deakin Crick (2003), and Marshall and Drummond (2006) confirmed this idea, discovering that a strong emphasis on summative assessment may have an adverse impact on students' willingness to learn. In practice, a focus on summative assessment often results in teaching to the test and, therefore, away from deep learning (James, 2006; Lau, 2016). Additionally, in a study that compared the value of information collected from both types of assessment, Andrade and Cizek (2010) found that summative assessment information provided little benefit to the procedures necessary to avoid surface learning. A traditional concentration on summative assessment marks a need for a

"paradigm shift", as Gipps (1994) argued even before the turn of the century.

The traditional view, as outlined by Gipps (1994), is that assessment should be a norm-referenced judgement, with all individuals' examinations being conducted, marked, and interpreted in the same way to enable cross-comparison of results. Consequently, as Gipps (1994) related, exams and standardised assessments are typical in the traditional view of assessment. Moreover, Gipps (1994) argued that the traditional view operates on a belief in the unchanging and inherent character of intelligence. Consequently, judgement in the traditional view focuses on classification, categorisation, and selection. In contrast, the criterion-referencing view characterises learning outcomes according to the definite assessment standards adopted by teachers instead of judging students against their fellow learners (Gipps, 1994; Taylor, 1994). This new view of assessment, as Gipps (1994) explained, focuses on assessing educational experiences' processes and outcomes.

However, even from a traditional assessment perspective, both types of assessment are critical to the learning process; and, according to Broadfoot and Black (2004), the optimal environment for learning is one in which formative and summative assessment methods are balanced. Black and Wiliam (1998) suggested that formative and summative assessment are primarily connected; they support each other. Likewise, Biggs (1998) argued that formative and summative assessment should not be seen as "two different trees" but as "the backside of an elephant" (p. 108). Taras (2007) clarified this by explaining that "each limb must work with the other in order for the whole to work; the animal is stronger as it is better balanced, and without one back leg the elephant would fall over" (p. 64). This relates to the sense that an elephant would keel over if it had a hind leg missing, as more effective balancing and strength is achieved by the elephant when its legs are working in coordination and harmony, enabling the entire body to move effectively. This metaphor indicates the importance of Bloom et al.'s (1971) initial idea that both kinds of assessment need to be linked and work in harmony. Indeed, different types of assessment should be interpreted in light of the extent to which the information produced may be used to guide future learning; rather than simply categorized them as summative or formative (Black et al., 2011). Consequently, instead of pushing for summative assessment to be abandoned as argued by Gipps (1994), supporting learning could be achieved through exploring and understanding the various purposes for various approaches of assessment and how various assessment approaches may operate in coherence to enhance learning.

The three main approaches to assessment are Assessment of Learning (AoL), Assessment for Learning (AfL), and Assessment as Learning (AaL). As the terms indicate, each has a distinct and unique character and function and, therefore, serves a precise purpose. Notably, summative assessment and formative assessment are sometimes referred to as AoL and AfL, respectively.

Within the literature, AoL is sometimes referred to as traditional assessment (Earl, 2003/2013). Earl (2013) stated, for example, that AoL is a traditional approach with a prime concentration on the production and gathering of evidence or information for a summative judgement of students' performance. In AoL, learners are usually passive subjects who take tests designed to evaluate and measure progress and performance; these tests are not administered during the learning/teaching process, but at the end of a teaching unit in order to measure students' knowledge of decontextualised content (Harlen & James, 1997). In short, AoL concentrates on summative purposes.

AfL, in comparison, denotes a substantial change in the assessment's role in the learning process. Growing interest in AfL, in turn, suggests a shift in concentration from improving students' scores to improving students' learning (Murtagh & Webster, 2010). A substantial movement has emerged in the assessment culture of higher education sectors. Namely, AfL is being promoted to provide further information to enhance the educational process for both teachers and learners and decrease attention on judgment, classification, and categorisation (Kennedy et al., 2008).

AfL concepts and practices have been promoted by the Assessment Reform Group that was formed in 1989 by a group of educational assessment researchers with the support of the British Educational Research Association. In 2002, the Assessment Reform Group found that assessment which promotes learning has the following characteristics: sharing learning objectives with learners, informing learners of the criteria they are aiming for, engaging learners in self-assessment and peer-assessment, providing feedback for the learners in order to recognise and plan for their next steps, having the confidence of learners' ability to improve, and ensuring the reflect on assessment information from both teachers and students. The purpose of these characteristics is helping to decrease the difference between the actual level of learners and the intended objectives, and enhancing learners' ability to self-assess and monitor their learning and progress. Based on these characteristics, the group used the term *AfL*, defining it as the process of searching for and construing evidence that can be applied by teachers and students alike in order to

determine not only the stage of learning that students have achieved but also what students' future direction should be and the most effective means of arriving there (ARG, 2002). Consequently, AfL concentrates on the process of learning and supporting students to address their individual learning gaps (Sadler, 1989). These characteristics, when used efficiently, can improve learning and teaching.

Once formative assessment is effectively applied, learning occurs in Vygotsky's Zone of Proximal Development (ZPD), also known as Zone of Potential Development (Vygotsky & Cole, 1978). In this zone, teachers are responsible for structuring and leading learning by way of scaffolding information with the students based on a process of gathering and interpreting evidence (Vygotsky & Cole, 1978). Thus, through applying formative assessment, instructors and learners are motivated to work on a consistent basis in the ZDP. AfL in the zone of proximal development helps teachers to provide experiences and support to their students, enabling students to gain new knowledge and understandings. Learners may also learn and benefit from classroom interaction and collaboration.

However, the effectiveness of formative assessment may be undermined when there is significant pressure placed on teachers to improve test performance, as is the case with many accountability systems (Bennett, 2011; Torrance, 2012). AfL and formative assessment were criticised by Torrance (2012) and Bennett (2011), who argued that use of formative assessment is being reduced to the monitoring of students' achievements through regular tests or other techniques used to "teach to the test". Their criticism of AfL can be linked to the work of Marshall and Drummond (2006), which found that some AfL practices represent only the "letter" of AfL, not the "spirit" of AfL; practices that do capture the spirit of AfL are discussed later in Section 3.3.3.2.

Within an AfL context, though both teachers and students are central in the process, the teacher's role remains more significant. Earl (2003/2013) thus proposed the concept of AaL to complement, support, and extend AfL, as AaL focuses on the students and their role in learning, while AfL keeps teachers centred.

AaL can be seen as an extension of AfL, as both approaches focus on formative assessment to promote learning (Berry, 2008). Dann (2002/2012) first introduced AaL as "a process through which pupil involvement in assessment can feature as part of learning" (p. 153). Earl and Katz (2006) then highlighted AaL as a metacognitive process in which learners are encouraged to monitor and employ self-regulatory skills in their thinking processes,

with an emphasis on the value and importance of encouraging learners to be their own assessors. AaL aims to provide learners with valuable opportunities to think, self-assess, self-evaluate, self-reflect, and self-correct both their learning progress and the strategies they implement. It thereby aims to help learners to identify their strengths and weaknesses, so learners can improve their own learning (Lee & Mak, 2014). Although learners are the primary focus of AaL, teachers are involved in AaL through their effort to serve students' learning by maximising learner involvement and enhancing learner motivation; teachers can improve the attention, engagement, and productivity of their students' learning by enhancing the effectiveness of their instruction through making it more goal-oriented, student-centred, and effective (Earl, 2013). In AaL, in short, the instructor serves as a facilitator, assisting students in their learning, rather than the key source of knowledge and information.

Earl and Katz (2006) argued that each of the three approaches is useful and serves a distinct function. Mok (2012) highlighted the need to find a balance between assessment as, of, and for learning in order to gain the educational and learning advantages of each approach. However, achieving a balance among the three approaches is not always easy. The ongoing attempts for balancing and combining these approaches of assessment over the years have led to the emergence of learning-oriented assessment (LOA), a concept originated by Carless et al. (2006). As the term indicates, LOA is a type of assessment in which the major concentration is on the possibility for students to establish productive learning processes (Carless, 2007, 2014). Carless (2007, 2014) highlighted that both formative and summative assessments have the potential to enhance learning as long as the primary emphasis is on creating appropriate student learning opportunities. Carless et al. (2006), Carless (2007, 2014), and Keppell et al. (2006) thus suggested that LOA should aim to stress the educational aspects of assessment and to address the question of how assessment may be utilised more effectively to improve students' learning. In agreement with Carless, Purpura and Turner (2014) emphasised that learning should be central to the curriculum and should serve as the guiding principle for instructional decision-making; specifically, they argued that learning could be prompted by LOA through situating assessment as a vital component of learning. LOA is a relatively new approach to assessment that has the potential to significantly increase students' learning.

Moreover, Carless (2007, 2014) provided a framework to understand Learning-Oriented Assessment (LOA) that is summarised in the following principles:

- Principle 1: Assessment tasks should be designed to stimulate sound learning practices amongst students.
- Principle 2: Assessment should involve students actively in engaging with criteria, quality, and their own and/or peers' performance.
- Principle 3: Feedback should be timely and forward-looking so as to support current and future student learning.

(Carless, 2007, p. 59)

The first principle suggests that including assessment activities that represent the intended learning goals can prep students for deep learning experiences (Carless, 2007, 2014). The second principle suggests that the participation of students in assessment enables them to gain deeper knowledge of learning objectives and to engage more actively with criteria and standards (Carless, 2007, 2014). The third and last principle suggests that, for an assessment task to be effective in promoting learning, the task should give students relevant and appropriate feedback that the students can feedforward and incorporate into future work (Carless, 2007, 2014). Based on these principles, Purpura (2016) described LOA as a cognitive, collaborative, and learner-centred approach to learning. Consequently, it could be argued that LOA can boost students' self-regulated learning and autonomy through creating an active learning environment which concentrates on the use of metacognitive strategies and feedback.

Furthermore, researchers including Carless et al. (2006), Carless (2007, 2014), Willey and Gardner (2009), and Purpura and Turner (2014) have suggested that Learning-oriented assessment could be developed through self-assessment. Engaging students in self-assessment may help students to develop learner autonomy, self-regulatory skills, and critical thinking. In turn, students can focus on meeting learning objectives. Evidence has revealed that learners that engaged in self- and peer-assessment may increase their likelihood of comprehending and engaging with the learning objectives, which has an impact on their performance (Daley & Nisa, 2016; Dann, 2012; Klenowski, 2009). Evans (2013) review of literature on self-assessment highlighted that self-assessment enhances learners' confidence and empowers them to take charge of their own learning process. Self-assessment is an influential approach, particularly in the EFL context. Such approaches may lead to the enhancement of learner's confidence, independence, communication skills, and deepened learning, as well as to critical thinking skills' developments. Hence, self-assessment remains significant owing to its positive influence on students' learning skills

and abilities. Thus, the following section explores self-assessment in the EFL context to understand its role in lifelong learning and in the promotion of self-regulated learning and critical thinking in this specific context.

3.3.2. Self-assessment in the EFL context

During the 1970s and 1980s, very limited attention was devoted to self-assessment within the higher education context, particularly in the field of second and foreign language learning (Coombe & Canning, 2002). However, since the 1990s, the increasing prominence of student-focused learning has led to self-assessment gaining ground as a means of assessment (Coombe & Canning, 2002). Thus, according to Harris (1997), learners are generally now required to assume a degree of responsibility for the planning, organisation, implementation, and evaluation of their learning. Traditionally, the obligation to regulate these four components was assigned to teaching staff alone. Other trends in pedagogy include the increasing relevance of lifelong learning and reflective learning in higher education (Sambell et al., 2012); the value conferred on learner autonomy (Hunt et al., 1989); and the increased focus on needs analysis (Blanche & Merino, 1989), All have led self-assessment to become accepted as a valuable means of assessment.

3.3.2.1. Definition of self-assessment

The concept of self-assessment is open to dispute, with no agreed definition of the term. The developing concept draws on various perspectives of different writers and researchers. For example, according to Dickinson (1987), self-assessment denotes a method through which students can assess and track their levels of expertise, performance, and comprehension in order to obtain an overall of their academic progress. Such a characterisation places self-assessment in a constructivist framework. In contrast, Underhill (1991) defined self-assessment as a form of testing that can encourage the participation of learners in the evaluation of their expressive or oral skills; Underhill noted self-assessment is, therefore, the easiest, cheapest, and quickest form of assessment. Underhill's definition aligns self-assessment with behaviourist theories. The definition adopted by Blatchford (1997) adds a comparative element to self-assessment. Namely, Blatchford defines it as the process by which learners can measure and judge their own improvement in relation to the performance of their peers. However, others have dismissed the notion that self-assessment is in any way comparative, insisting that it only involves students making judgements about their own achievements or failings in a way that serves

to expand self-awareness. For example, Dikel (2005) contended that self-assessment comprises a route by which students arrive at an awareness of their personal learning traits, their personal learning preferences, and their responses to particular learning scenarios. Similar to Dikel's definition, Montgomery (2001) definition describes it as a means of judging, by which students can judge their own performance and learning journey. Specifically, as Mousavi (1999) stated, self-assessment is a learner's personal evaluation of their linguistic competence and their ability to employ the target language in multiple scenarios. Nevertheless, most definitions of self-assessment characterise self-assessment as one's abilities, processes, or products, that can add to the complexity of defining the term. For example, Epstein et al. (2008) defined self-assessment as a continuous process of self-monitoring from one point in time to the next; it refers to observing our individual actions, to curiosity about the implications of those actions, and to our willingness to develop our behaviour and thinking through learning from our observations. Brown and Harris (2013) defined it as a descriptive and evaluative act undertaken by a learner in relation to their own academic abilities and progress. Panadero, Brown, et al. (2016) described it as a "wide variety of mechanisms and techniques through which students describe (i.e. assess) and possibly assign merit or worth to (i.e. evaluate) the qualities of their own learning processes and products" (p. 804).

The multiplicity of definitions for the concept of self-assessment also demonstrates a lack of agreement regarding the term itself. Self-assessment may be referred to as self-evaluation, self-rating, or self-appraisal. These terms tend to be variously applied to denote self-assessment and sometimes are used interchangeably within the literature. Klenowski (1995) argued that self-evaluation refers to the assessment or appraisal of the value of performance and the recognition of positive and negative aspects with the intention of enhancing future learning outcomes. Klenowski's definition is somewhat broader than most self-assessment definitions, which are more restricted, focused instead on the processes in which learners engage in order to determine the grade to which they are entitled. Klenowski's definition could be seen as broadly parallel to that employed by Dickinson (1987) in that both definitions are contextualised within a constructivist frame, which emphasises evaluation and information gathering on areas that require improving. The term *self-rating* tends to refer to learners using a scale to score their progress and performance (Behar-Horenstein et al., 2018; Cheng et al., 2010; Goh et al., 2010; Taheri et al., 2014; Williamson, 2007). Whilst the terms *self-assessment* and *self-appraisal* are often used interchangeably (McLeod, 1997; Sobral, 2004), self-appraisal is also used to indicate self-evaluation of success by certain writers such as Van Praag et al. (2017).

Due to the difficulties in defining self-assessment, some researchers, including Haughton and Dickinson (1988), Oscarson (1989), Bachman (2000), have endeavoured to define the term based on the purpose of self-assessment. Along that same line, Andrade and Cizek (2010) and Andrade (2018) argued that most definitions are missing the purpose of the act of self-assessment. Within the literature, formative and summative assessment are the two identified types of self-assessment (Andrade & Cizek, 2010; Andrade, 2018; Andrade & Brown, 2016; Bachman, 2000; Barnes, 2006; Brown et al., 2015; Carless, 2005; Dickinson, 1987; Oscarson, 1989).

Andrade and Cizek (2010), Andrade (2018), and Andrade (2019) argued that "self-assessment is feedback" and that the drive of feedback is to guide alterations that enable deep and enhanced learning and achievement; therefore, self-assessment is meant to generate feedback with the purpose of promoting learning and improving performance. This argument emphasises feedback as a vital part of the student process of self-assessment, and it could help avoid superficial implementations of self-assessment, as discussed in the following section.

3.3.2.2. Self-assessment from "letter" to the "spirit"

Before discussing the implementation of self-assessment within the context of EFL learning, it is important to clarify that self-assessment is far more than simply applying a technique. Superficial implementations of self-assessment must be avoided to ensure its benefits. Self-assessment as part of AfL or AaL is grounded in a constructivist approach (Marshall & Drummond, 2006; Munns & Woodward, 2006; Murphy, 2008). As noted, AaL is derived from AfL. The constructivist approach of learning is a theory developed by Piaget (1978) that concentrates on building knowledge through experiences' interpretations. Through scaffolding, students can build new knowledge grounded on previous knowledge and experience (Hmelo-Silver et al., 2007). However, AfL and AaL could be seen as forms of interaction between teachers and students or between students. Assessment is a complex cultural process, wherein the connections between learners, teachers, and tasks are explored in their context (Elwood, 2006). Accordingly, the constructivist approach to self-assessment is questioned so that self-assessment can be seen from a sociocultural perspective. The sociocultural approach enables teachers to navigate and understand the complexities of their context in promoting learner autonomy (Willis, 2009). It recognises that "activities do not exist in isolation rather they are part of broader systems of relations, social structures, in which they have meaning" (Murphy et al., 2006).

Social and cultural contexts have a powerful impact on classroom assessment and learners' autonomy, and, without comprehension of their impact, AfL may well be "part of the futile search for a universal, culture-free, 'teacher-proof' approach to education" (Wells & Claxton, 2002, p. 6). In this regard, Marshall and Drummond (2006) differentiated between the "letter" and "spirit" of AfL; namely, they indicated that teachers that understand the importance of social contact and the importance of sharing the learning responsibility along with their learners, will eventually lead to a positive influence on students' performance and autonomy with the support of AfL practices. The spirit of AfL is present in teaching when the environment of the classroom supports socially constructed learning and when the goal is to increase learner autonomy. Therefore, to avoid superficial implementation of self-assessment, and to understand AfL with sociocultural theory, the origin of the theory including the work of Dewey (1910/1997) and Vygotsky (1978), is explored in the following section.

3.3.2.3. Origin of sociocultural theory

Sociocultural theory can be linked to the work of John Dewey (1910/1997) and Lev Vygotsky (1978), two prominent psychologists in developing the sociocultural theories of learning. Glassman (2001) argued that Dewey's work may possibly have had an impact on Vygotsky's thinking. Though both psychologists did differ on some substantial aspects of their comprehension of learning, both had primarily the same viewpoint regarding the socially constructed aspect of learning and the concept of development toward learner's autonomy (Glassman, 2001). Dewey (1910/1997) argued that the teacher–student relationship facilitates learning, such that understanding is developed through social contact and "scientific inquiry" instead of the traditional instruction of a subject. Learning processes should thus involve students in active learning that incorporates teacher-student interaction in a social context. Moreover, Vygotsky (1978) concentrated on the function of language as a cultural instrument rather than on individual development as an end goal, by including children in activities and allowing them to engage with more experienced individuals, and thereby foster stronger social cohesion (Vygotsky & Cole, 1978). Vygotsky (1978) further argued that, when adults use language to communicate as an attempt to complete a task, language is viewed as a tool to facilitates activity. Viewing language in this way facilitates the internalisation of learning and gives a foundation for comprehension of their following activity.

Vygotsky (1978), in explaining his concept of the zone of proximal development (ZDP), emphasised the gaps between degrees of possible development; those gaps, he argued, are influenced by "independent problem solving and the level of potential development as determined through problem solving under adult guidance, or in collaboration with more capable peers" (Vygotsky & Cole, 1978, p. 86). Vygotsky (1978) highlighted the significance of the instructor-student relationship, wherein the instructor recognises and scaffolds tasks within a ZDP that is slightly above the developmental level of the learner (Vygotsky & Cole, 1978). This approach emphasises that the teacher's role is guaranteeing that all learners may collaborate and work in a social setting to build new knowledge and understanding, thereby fostering and encouraging their individual learning.

Drawing on Vygotsky' concept of ZPD, Gipps (2002) uses the notion to support her position on the transition from AfL. The zone of proximal development is the focal point of Gipps (2002) argument on assessment, in that it both reflects and promotes the learning process of the students. Gipps (2002) outlines four critical characteristics of assessment:

> The critical role of tools in human activity and implications of offering assistance and guidance during the course of an assessment... The inseparability of the social, affective and cognitive dimensions of action and interaction and hence the implication that learners should be assessed not in isolation and in competition but in groups and social settings... The relationship between expert and apprentice...and the implications of this for the assessment relationship... (p. 74)

Self-assessment, within the context of the sociocultural, can be seen as a cultural and dialectical process that allows learners to improve their knowledge and have control over their learning process.

This assists in comprehending self-assessment in a sociocultural context; as an interactional process, students and teachers or peers collaborate with the purpose of improving learning. This suggests that, through interaction and cooperation, learners may explore deeply and gain knowledge from one another.

3.3.2.4. Self-assessment and self-regulated learning from a sociocultural perspective

A sociocultural perspective casts doubt on the more traditional notion of learner autonomy as "the capacity for detachment, critical reflection, decision-making, and independent action" (Little, 1991, p. 4). From a sociocultural perspective, learning is situated within a

cohesive cultural process; the social context and the learner are interrelated in a manner that influences and shapes the learning process, with the social environment being influenced by and influencing the student (James, 2006). The learner acquires competence through language use and participation in social interactions, which include collaboration with the instructor. Willis (2009) highlighted that learner autonomy is characterised as the socially created identity of a self-monitoring learner that engages in culturally acceptable behaviours in a community of practice. Willis's (2009) interpretation emphasises the view that students develop learner autonomy in a social setting rather than in solitude. Willis's definition of learner autonomy is influenced by Lave and Wenger (1991) definition of identities as "long-term, living relationships between persons and their place and participation in communities of practice" (p. 93). In Willis's (2009) definition of learner autonomy, self-monitoring entails the ability to comprehend and control learning at the time it occurs, with dependence on the teacher reducing over time. Regarding this definition of learner autonomy, Ratner (2000) further argued that the learner's autonomy transitions from a stable to a prospective state, and only with social interaction can the role of an autonomous learner be fulfilled. The learner is thus actively engaged in negotiating their own learning identity.

Within a sociocultural perspective, to assist learners in developing self-regulatory skills, both teacher- and student-led AfL and AaL activities should be made more interactive and dialogic (Nicol, 2010). Making these practices more interactive will entail making self-assessment dialogic as well in order to develop students into autonomous and self-regulating learners. Nicol and Macfarlane-Dick (2006) and Nicol (2010) argued that feedback should be considered a two-way, dialogical process including teacher-student and/or peer-peer interaction, as well as the learner's active involvement through self-feedback. Interpretations are exchanged, meanings are negotiated, and expectations are defined in such an engaged discussion; this approach contrasts with conventional assessment practices that follow an old paradigm centred on one-way information flow from the teacher to the student (Carless, 2012). Feedback as dialogue is argued to promote learner' active involvement with feedback (Price et al., 2011), and self-regulated learning (Winstone et al., 2017). With self-assessment, students thus play an active role in providing feedback through self-assessment or peer-assessment to develop their performance. Students can learn to accept and comprehend useful and meaningful comments in an approachable way. Panadero, Brown, et al. (2016) argued that, through teacher and peer evaluation or feedback, novice self-assessors could become aware of the possibility of inaccuracy in their assessment. Additionally, it is argued in the work of

Panadero, Brown, et al. (2016), Panadero and Alonso-Tapia (2013), and Cao and Nietfeld (2005) that providing feedback to students is necessary for them to develop into more accurate self-assessors, and without feedback self-assessment appears to be highly dependent on individual qualities and differences.

Moreover, through self-assessment, students can also practise detached judgement, learning to become more critical of their own work and, thus, improving their self-regulatory skills. The next section discusses the concepts of self-regulated learning and its compatibility with self-assessment.

3.3.3. Self-assessment and areas of self-regulated learning

Self-regulated learning provides a theoretical framework to the notion of self-assessment by emphasising the learners' active role in the learning process (Panadero, Brown, et al., 2016). It is argued in the literature that engaging students in self-assessment could develop their self-regulated learning skills (Andrade & Cizek, 2010; Andrade, 2019; Panadero & Alonso-Tapia, 2013; Panadero et al., 2018; Panadero, Jonsson, et al., 2016). Additionally, several research studies as those conducted by Yan (2020), Brown and Harris (2013), and Panadero et al. (2018) have demonstrated that self-assessment is a critical component of SRL and occurs throughout each phase in the SRL process. Recently, Panadero et al. (2017) meta-analytic review and Andrade (2019) critical review on the relationship between self-assessment and self-regulated learning suggested that self-assessment interventions are beneficial for students' self-regulated learning skills. By placing self-assessment within the theoretical framework of self-regulated learning, which includes cognitive, metacognitive, self-efficacy and motivation theories, the following section explores the connections between the two concepts and highlights the importance of self-assessment to promoting students' self-regulatory skills.

3.3.3.1. Self-assessment and achievement goal theory

Achievement goal theory is a cognitive theory concerned with both the manner in which learners internalise diverse ability objectives and the impact of these objectives on self-assessment, learner persistence, and learning outcomes (Senko et al., 2011). Studies within the field have tended to emphasis two types of goals: mastery goals and performance goals (see, for example, Harackiewicz et al. (2002); Maehr and Zusho (2009); Pintrich et al. (2003); Senko et al. (2011)). Mastery goals push the learner to concentrate on a given task,

specifically on the knowledge and skills required to complete it (Maehr & Zusho, 2009). Mastery is acquired through cognitive processes, including self-monitoring, thinking, and problem-solving (Maehr & Zusho, 2009). When working towards mastery goals, learners tend to completely involve themselves in tasks and activities, monitoring their own progress throughout (Pintrich et al., 2003). Conversely, performance goals emphasise outcomes and results, including grades and scores (Senko et al., 2011). Performance goals emphasise focus on the outcome and identifying ways to ensure the outcome; more emphasis is placed on ensuring the grade than on fostering a deeper knowledge. Performance goals can, therefore, prompt learners to develop negative perceptions regarding their abilities and performance. Performance goals may also serve to reinforce compliance and conformity, which then prevent any real improvement in comprehension or proficiency since the grade or score becomes more important than authentic learning.

The presence of performance goals implies that the process of monitoring and evaluating learning is an external one. Conversely, mastery goals suggest that the monitoring and evaluation are, at least to a degree, internal. Self-assessment is invaluable in the context of mastery goals, since it facilitates self-awareness in learners regarding their knowledge and skill levels. However, performance orientation is dependent on the teachers and other staff who arrange learning activities, define success and failure, and assess outcomes. Thus, augmenting skills in relation to self-assessment improves mastery orientation.

3.3.3.2. Self-assessment and metacognition

Metacognition is the process of "thinking about thinking" (Flavell, 1979, p. 906). Unsurprisingly, it has a marked influence on self-assessment. Metacognitive and self-regulated learning strategies are fundamental for learner success (Bol & Garner, 2011). Self-regulated learning is an active, practical process in which learners determine goals for their learning; these goals, along with the contextual characteristics of the learning environment, will guide learners' monitoring and regulating of their own cognition, motivation, and behaviour (Pintrich, 2000). To become active in their learning process, students should choose and adapt their learning strategies and reflect on the effectiveness of those learning strategies (Cleary & Zimmerman, 2012). Metacognitive strategies include planning, monitoring, and evaluating learning performance (Pintrich, 2000). These strategies can be regarded as the means by which learners track, organise, and arrange their learning. Metacognitive strategies also include learners developing awareness of what activities and techniques they employ during learning (Bol & Garner, 2011). Thus, through

the process of metacognition, learners can acquire the ability to understand which learning strategies have the most effective outcomes in specific learning contexts. Further, learners can realise what measures must be undertaken to enhance outcomes. However, to conduct metacognition effectively, students must be aware of their work and be able to monitor, evaluate, and realise what to do to improve performance; these capacities require skills such as checking, understanding, predicting outcomes, planning activities, managing time, and switching to different learning activities as needed (Berry, 2013).

Metacognitive strategies arguably require self-regulatory skills on the part of the learner, and they certainly have positive links to self-assessment and enhanced learning performance. Learners must be able to monitor, evaluate, and assess their own learning both during and following the learning experience. In this way, learners become engaged participants, utilising any information produced via assessment and assuming responsibility for their future learning performance (Berry, 2013). Moreover, as Berry (2013) noted, the intention embodied within metacognition-related assessment is to facilitate the development of students as independent learners. However, metacognition-related assessment necessitates learner awareness in relation to what is expected from them, as well as the ability to regulate and evaluate their own progress. Using the information acquired through self-assessment, learners are in a position to manage their learning in a way which allows the prompt achieving of their objectives. Clearly, this conceptualisation of assessment places the active role of learners at its core.

3.3.3.3. Self-assessment, self-efficacy, and motivation

Conceptually, self-efficacy is closely related to the notion of learner autonomy, which exists in the discussion of language learning under the term *self-regulation*. As Dann (2012) noted, learner self-efficacy determines how a learner proceeds in any given teaching, learning, or assessment scenario. Hence, beliefs concerning self-efficacy have a causal relationship with a learner's employment of self-regulatory behaviours, such as time management, learning strategies, and the establishment of personal academic goals (Dann, 2012). In other words, the greater the academic capacity students perceive themselves to possess, the more advanced the targets they establish for themselves. Should students fail to achieve their objectives, the more self-efficacious amongst them will simply redouble their efforts, whilst the less self-efficacious will tend to revoke their active pursuit of academic success (Zimmerman, 2000b). Thus, a learner's self-belief regarding their

academic potential or prowess has a direct and marked influence on their capacity to realise their goals (Zimmerman, 2000b).

Self-regulation comprises three discrete stages: forethought, performance, and self-reflection (Cleary & Zimmerman, 2012; Zimmerman, 2000a; Zimmerman & Cleary, 2009). Self-efficacy is a sub-constituent of self-regulation, belonging specifically to the forethought phase, which encompasses notions existing prior to the learning performance (Cleary & Zimmerman, 2012; Zimmerman & Cleary, 2009). Making estimations pertaining to potential performance and ability, and to the probability that effort will be met with success are crucial aspects of self-efficacy (Cleary & Zimmerman, 2012; Zimmerman & Cleary, 2009). Such self-perception emerges alongside the capacity of learners to incrementally link their successes and failures to specific causal factors. Situation and context are crucial to self-efficacy. Moreover, self-perception is vital because it correlates with learners' beliefs concerning their general capabilities, knowledge base, and skills (Zimmerman, 2000a, 2000b). Thus, learners who possess high expectations in relation to their potential are more likely to act in a manner indicative of persistence. Conversely, learners with lower expectations regarding their abilities are less prone to persevere in learning and may demonstrably shun certain tasks or relinquish their studies altogether. Therefore, in the opinion of McCarthy et al. (1985), the employment of self-assessment techniques presents itself as a valuable way to enhance learners' self-efficacy and, consequently, their abilities and performance levels.

Self-assessment must be undertaken for a number of reasons. Students need to self-assess in order to identify when they are learning, how much work is necessary for achievement, when they have achieved success, when they have made a mistake, and which learning methods work best for them (McMillan & Hearn, 2008). Effective and honest self-evaluation facilitates the identification of what has been successfully grasped versus what is yet to be mastered. Learners who succeed in performing reasonably demanding tasks can tribute their achievement to their skill and effort, not external variables such as peer support. Identifying internal strengths depends on the capacity to make accurate self-assessments and evaluations.

However, learners must not merely possess convictions concerning self-efficacy (Zimmerman, 2000a). Learners should also be stimulated to become active learners. In order to develop as active learners, students must be autonomously inspired to take all they can from their learning environment. Learners should formulate judgements and make

decisions pertaining to the degree to which they wish to actively learn (Dann, 2012). Self-beliefs regarding personal learning capacities will clearly inform this effort to some extent. According to Rheinberg et al. (2000), learning-focused motivation encompasses multiple traits, including motives, objectives, intentions, predilections, and situational factors such as potential advantages and disadvantages of tasks and the nature of activities involved.

Thus, self-assessment is a principal classroom-based activity which can be invoked to teach learners about learning. Moreover, as Dann (2012) maintained, self-assessment can educate learners on how to review their own work frankly, impartially, and constructively. It can also teach learners to express opinions about their personal learning experiences, to celebrate success, and to formulate realistic goals. Furthermore, the capacity of self-assessment to determine self-efficacy and other aspects of self-regulation should not be forgotten (Zimmerman, 1989). Thus, when learners possess negative attitudes towards their personal learning, self-assessment is a valuable strategy which can be employed to manifest and ground more positive aspects of the process. Before exploring the different self-assessment strategies used in the EFL context, it is necessary to explore the ideas around speaking and assessment of speaking in the EFL context in the following section, as this thesis concentrates on speaking classroom.

3.3.4. Speaking and its assessment

Within an EFL context, speaking skills are a vital component of the curriculum (Luoma, 2004). Speaking skills are fundamental for global mobility, admission into higher education, and employment in the globalised society (Fulcher, 2015; Isaacs & Isaacs, 2016). Speaking, according to Bailey (2003), is the systematic production of utterances to communicate meaning, namely an oral, aural, and productive skill that occurs in real time. To illustrate, as speech is generated orally, speaking is oral; as the response is frequently related to input audibly received, speaking is also aural; as language is directed outward, speaking is productive too; and lastly, as the other speaker should wait for the person who is speaking to finish speaking before they can speak, and as the speaker is unable to revise his response as he could in writing, speaking occurs in real time (Bailey, 2003). Notwithstanding, regardless of the prominence of speaking in language assessment and pedagogy, several researchers, including Correia (2016); Fan and Yan (2020); Guettal (2008); Levelt (1994); Rychtarik (2014); Schmidt (1992), have viewed speaking as an intangible construct that can be difficult to conceptualise and assess. Hughes (2003) indicated that selecting the appropriate assessment is a major challenge in assessing

speaking. Along with selection of the appropriate assessment, O'malley and Pierce (1996) suggested that identification of the assessment criteria is another major challenge in assessing speaking. These challenges in conceptualising and assessing speaking, according to Luoma (2004) and Carter and McCarthy (2017), could first probably be due to the dynamic and context-specific character of spoken language. Second, Luoma (2004) and Carter and McCarthy (2017) have also argued that the challenges could also be due to the numerous forms that spoken language can take, such as paired conversation, group discussion, and monologs, and to the various situations in which speech occurs, such as intentional or unplanned situations.

In an attempt to help determine what to assess when it comes to speaking, Brown and Abeywickrama (2004) classified speaking skills into micro-skills and macro-skills. Micro-skills encompass the ability to produce small units of language, including words, morphemes, phonemes, collocations, and phrase units, while macro-skills concern the larger components of speaking, including style, fluency, function, cohesion, discourse nonverbal communication, and strategic options (Brown & Abeywickrama, 2004) (see Table 3-3 for a list of micro- and macro-skills of speaking).

Table 3-3 a list of micro- and macro-skills of speaking (Brown & Abeywickrama, 2004)

Micro-skills	Macro-skills
• Producing differences among English phonemes and allophonic variants. • Producing chunks of language of different lengths. • Producing English stress patterns, words in stressed and unstressed positions, rhythmic structure, and intonation contours. • Producing reduced forms of words and phrases. • Using an adequate number of lexical units (words) to accomplish pragmatic purposes. • Producing fluent speech at different rates of delivery. • Monitoring one's own oral production and using various strategic devices – pauses, fillers, self-corrections, backtracking – to enhance the clarity of the message. • Using grammatical word class (nouns, verbs, etc.) systems (tense, agreement, pluralisation); word order; patterns; rules; and elliptical forms. • Producing speech in natural constituents: in appropriate phrases, pause groups, breathe groups, and sentence constituents. • Expressing a particular meaning in different grammatical forms. • Using cohesive devices in spoken discourse.	• Appropriately accomplishing communicative functions according to situations, participants, and goals. • Using appropriate styles, registers, implicature, redundancies, pragmatic conventions, conversation rules, floor keeping and yielding, interrupting, and other sociolinguistic features in face-to-face conversations. • Conveying links and connections between events and communicating such relations as focal and peripheral ideas, events and feelings, new information and given information, and generalization and exemplification. • Conveying facial features, kinesics, body language, and other nonverbal cues along with verbal language. • Developing and using a battery of speaking strategies, such as emphasizing key words, rephrasing, providing a context for interpreting the meaning of words, appealing for help, and accurately assessing how well your interlocutor is understanding you.

Moreover, building on the work of Heaton (1988), Hughes (2003), and O'malley and Pierce (1996), Brown and Abeywickrama (2004) identified five categories of speaking assessment practices:

1- Imitative practices, or repeating a small stretch of language to assess competence in phonetics, prosody, lexicon, and grammar (pronunciation).
2- Intensive practices, such as via reading aloud (to assess stress-pattern, rhythm and pronunciation); directed response tasks (to assess a specific grammatical form or a transformation of a sentence); sentence/dialogue completion (to assess the micro-

skill of providing the right chunks of language and other pronunciation features); translation/interpreting games (to assess competence in conveying a message in the target language); and limited picture-cued tasks (to assess competence in explaining a plan, directions, and even opinions).

3- Responsive practices, or using small dialogues or responses to spoken prompts, such as via question and answer (to assess producing meaningful language in response); giving instruction and direction (to assess competence in describing a how-to description); and paraphrasing.
4- Interactive practices, or using larger dialogues in transactional and interactional conversation, such as via interview, drama-like tasks, discussions and conversations, and games.
5- Extensive practices, or monologues, such as eliciting speech through oral presentations or reports, picture-cued storytelling, retelling a story, and translation.

These categories of speaking assessment practices as classified by Brown and Abeywickrama (2004) emphasise the vital role of the purposes of assessment. Identifying and understanding the purpose of speaking assessment can contribute to the development of speaking tasks or assessments that target specific speaking abilities and achieve the intended objective, for example, to measure performance or support learning. Another starting point when it comes to speaking assessment is the language ability theories, such as Canale and Swain (1980) or Bachman (1990) theories, or language ability description, such as Common European Framework of Reference for Languages (CEFR) and the American Council on the Teaching of Foreign Languages Proficiency Guidelines (ACTFL). For example, Canale and Swain's (1980) theory of language ability focusses on four competencies: grammatical, strategic, discourse, and socio-linguistic competencies. On the same note, the CEFR's language ability description includes descriptions of the main four skills from beginning to advanced.

The above ideas related to speaking assessment could inform the creation of appropriate student learning opportunities through speaking self-assessment that stresses the educational aspects of assessment, promotes deep learning, and encourages active engagement in the learning process. The following section addresses the different self-assessment strategies used in the EFL context.

3.3.5. Self-assessment strategies in the EFL context

Self-assessment necessitates that learners employ a wide range of learning strategies and higher-order cognitive skills that offer more than just feedback it also offer guidance for future learning (Chamot & O'malley, 1994). Andrade (2019) reported in a recent critical review of research on learners' self-assessment that, without exception, the studies included in the review, which also included two meta-analyses (Graham et al., 2015; Sanchez et al., 2017), revealed a positive relationship between self-assessment and achievement. Research in the field of EFL learning in Saudi Arabia also revealed positive relationship between self-assessment and achievement in writing (Alshammari, 2016; Nalliveettil & Mahasneh, 2017) and reading (Rabiah, 2020); and general language improvement (Qasem, 2020).

Within the context of EFL, self-assessment may be incorporated through different strategies including learner diaries, dialogue journals, pre- and post-course writing, rating scales, checklists, and audio and video recording. The practical application of such strategies is clarified and explored in the following section.

3.3.5.1. Reflection pieces (learner diaries, dialogue journals, and reflection logs)

Diaries, dialogue journals, and reflection logs are types of reflective writing that Oskarsson (1984) and Dickinson (1987) initially proposed for students. Opp-Beckman and Klinghammer (2006) described the reflection log as written records by which personal actions, work, and attainment can be monitored by a student over the course of their learning. Actions that could be taken to achieve self-improvement, aspects that have been straightforward or challenging, issues that a student still has queries about, and critical reflection on the learning process could all be documented in a reflection log according to Opp-Beckman and Klinghammer (2006). This type of reflective writing can be appropriate at the level of a certain subject or an academic unit, as observed by North Carolina State Department (1999) (Jones et al., 1999). Additionally, Richards et al. (2002) noted that pupils' learning can also be monitored by teachers through reflection logs. Opp-Beckman and Klinghammer (2006) described dialogue journals as written reflections on the writer's own development and learning experiences generated on a day-to-day or week-by-week basis. Nunan et al. (1992) clarified that dialogue journals create teacher-student and student-student interaction outside classes.

Moreover, Boud and Falchikov (2007) pointed out that compared with a typical diary, the level of personal reflection will be less, although some critical reflection will be present in the journal. According to Boud and Falchikov (2007), diaries are effective means of assessment when the emphasis is on the learning process rather than the end result. Nunan et al. (1992) suggested that autonomous learning is promoted through diaries, as diaries encourage learners to take charge of their own learning. Furthermore, Nunan et al. (1992) suggested that the formulation of novel ideas, comprehension of challenging resources, improvement of confidence, and exchange of ideas with teachers are all possible subjects for learners to reflect on through diaries. Köller (2005) also noted that opportunities for critical reflection by learners are provided by learning diaries, which offer further autonomous learning opportunities. These methods offer a means for learners to methodically self-assess their progress over a length of time. Leaners are prompted to employ these journals or diaries for the free expression of elements of their learning; through these tools, learners can describe how they feel about their progress and explore their future intentions regarding the use of their newly acquired skills.

3.3.5.2. Audio-visual aids

Technology has enabled the recording of lessons, so audio-visual processes can be used to isolate and preserve certain elements in the process of learning. For instance, recording could be employed for self-assessment, wherein learners can use it to trace either themselves or their fellow students' performance. By listening to the recordings, students provide themselves with a valuable opportunity to gauge their own linguistic competence. When audio and visual recording is combined, students can evaluate their own paralinguistic skills, including tone and pitch of the voice, or body language as an adjunct to their language parallelism and linguistic performance. Repeated use of these techniques can allow learners to acquire a precise sense of their progress and performance, which is beneficial in terms of increasing learner confidence levels pertaining to oral productive skills. Audio recording appears to lend itself well to self-assessment in the context of speaking skills amongst language students in Saudi Arabia, a context in which obtaining permission to employ video recording may be fraught with obvious difficulties. VSR (video-stimulated recall) offers several benefits for independent language learners (Hurd & Lewis, 2008). It is capable of recording some aspects of classroom performance and permits the subsequent revisiting of data and consequent reflection. Moreover, decisions taken during lessons and emotions demonstrated at the time can be carefully reconsidered. There are notable drawbacks to VSR as well, though. Reitano (2006) noted, for example,

that VSR creates a predisposition to fixation upon physical appearance, can elicit embarrassment and self-consciousness, and may lead to intransigent attitudes regarding the objectivity of any observations.

Irrespective of affective factors, research by Ross (2006) concluded that, in combination with appropriate instruction, self-assessment by students constitutes a consistent and appropriate evaluation tool which can enhance both learning experiences and learning outcomes. According to Reitano (2006), VSR is the most efficacious way in which teaching staff can review their own teaching practices and continue their professional development. Thus, video or audio reflection has the potential to be used effectively in relation to language learners.

3.3.5.3. Self-assessment forms (rating scales/ criteria sheets/ rubrics and checklists)

Rating scales and checklists are common techniques in self-assessment. Such forms are frequently utilised to help language learners evaluate their apparent capabilities and linguistic proficiencies. Several research studies as those conducted by Coombe Coombe and Canning (2002); Coombe (1992) and Oskarsson (1984) have demonstrated the manner in which statements of ability, such as self-assessment forms, may be used by learners for a structured means of pondering and constructively assessing their own learning performance. A checklist is characterised by Herman (1992) as a set of dimensions, characteristics, or behaviours that that are ultimately rated with ‘no’ or ‘yes’ options. As Herman (1992) explained, a checklist assists the user to acknowledge the possession or lack of specific conduct or features; compared with rating scales, checklists often have more dimensions. Stimulating students’ active learning through reflecting on their work and helping them to learn in an effective way is perhaps the main use of checklists within the EFL learning context (Wragg, 2003). In contrast, rubrics offer guidance by providing criteria and descriptors of the different levels of proficiency or knowledge that could help learners while reflecting on their work (Griffith & Lim, 2012). It is argued by Andrade and Valtcheva (2009) and Panadero et al. (2013), that complex self-assessment scaffolding tools, such as rubrics, could assist beginners by providing clarity on what constitutes quality performance. The directions embodied within self-assessment forms and the criteria required can be either straightforward and brief, or extensive and complex. The extent of the measures employed in self-assessment forms typically correlates with the level of linguistic proficiency the students possess. However, the measures can also relate

to the specific issues which teaching staff wish to address and to any other language-related skills in which students must demonstrate ability (Diab, 2010a, 2010b; Harris, 1997; Kasule & Lunga, 2010; Oscarson, 2009).

Within the context of language learning, recent research indicates that rubrics can help students enhance performance and adopt self-regulatory learning strategies. For example, Baxa (2015), Herayati (2020), Ratminingsih et al. (2018), Weiss (2018), and Xu (2019) indicate that rubrics should clearly describe particular ideas for learners' work and influence learners' development through understanding the strengths and weaknesses of the learners. In the same vein, some research has found that the use of rubrics can improve performance and provide opportunities for self-regulatory learning use of strategy (Brown & Harris, 2013; Panadero & Alonso-Tapia, 2013; Taras, 2010). These findings imply that rubrics can be effective strategies that positively impact achievement and provide opportunities to use self-regulatory strategies. These tools enhance learners' awareness of the criteria that will likely be used to assess their performance, and Nielsen (2011) found that, without learners' comprehension and awareness of these criteria, assessment cannot be valid. The use of accurate, well-defined, and well-understood criteria and standards of performance, according to Panadero et al. (2013) and Panadero and Romero (2014), is essential for assisting students in assessing their own work to determine whether it meets standards and expectations, hence enhancing the accuracy of self-assessment.

3.3.6. Self-assessment as a means to enhance critical thinking along with self-regulated learning in the EFL context

Self-assessment can improve both self-regulated learning and critical thinking as it provides a link between critical thinking, self-regulated learning, and enhancing the learning (Watson, 2002). A student who is capable of thinking critically can monitor, analyse, judge, and select the best ways to learn; thus, a good critical thinker is a better self-assessor (Paul et al., 1997). In fact, self-assessment entails providing learners with the chance to monitor their development and solely concentrate on their learning (Harris, 1997). Learners should use critical thinking in order to make the most of this opportunity. Self-assessment enables learners to be actively involved in the process of assessment and to think deeply; it prompts students to develop cognitive skills such as critical thinking, collaboration, decision-making, self-monitoring and regulation, and problem solving; and it encourages them to criticise constructively, propose changes, reflect, and make sound decisions (Sluijsmans et al., 1998; Sung et al., 2005). Likewise, Smith (1997) stressed the

prominence of self-assessment and argued that self-assessment assists learners in becoming autonomous learners and that, in turn, critical thinking is one of the traits of autonomous learners. Moreover, Paul et al. (1997) proposed that self-assessment is a crucial factor for critical thinkers, who are capable of monitoring, analysing, judging, and selecting the most effective way of thinking; thus, a good critical thinker is a good self-assessor. Consequently, the argument is that self-regulated learning and critical thinking could be enhanced through providing learners with opportunities by which they can self-assess their performance as explained in detail in the above sections.

Reviewing the literature on self-assessment within EFL reveals that most studies investigating the impact of self-assessment have used observational approaches to characterise classroom activity. Few studies have offered evidence supporting their findings or illustrated the effectiveness of self-assessment in EFL classes. Nonetheless, in the last five years, there has been a recognised increase in the number of the studies on the implementation of self-assessment and its impact on achievement within the EFL context, e.g. Herayati (2020), Liu and Brantmeier (2019), Ratminingsih et al. (2018), Wang (2017), Elgadal (2017). Significantly, there is a noticeable difference between the number of the studies that have emerged in writing classrooms, on the use and effect of self-assessment in writing classrooms in comparison to the number of studies in speaking classrooms, especially in higher education. The review also revealed several recent empirical studies on the relationship between self-assessment and self-regulated learning in higher education within the EFL context, e.g., Papanthymou and Darra (2018), Duque Micán and Cuesta Medina (2017), Fathi et al. (2017), Dharma and Adiwijaya (2018). However, none of the studies have addressed to the issue in the Saudi context, particularly in English language courses in higher education. Additionally, neither the relationship between self-assessment and critical thinking nor the relationship between self-regulated learning and critical thinking have been sufficiently explored, especially in higher education within the Saudi context. Previous studies have emphasised the need for further research on the perception of students of the practices of assessment in EFL classrooms (Alhamami, 2019; Borg & Edmett, 2019; Kalra et al., 2017; Tong, 2011), which has not been sufficiently explored, especially in speaking classes. Such research could provide important evidence about the effectiveness of self-assessment. Learners' perceptions are important to understanding their preferences and likes and dislikes, or their desire to alter the learning/teaching process in the language classroom (Griffiths, 2007). Roghanizadeh (2011) highlighted that no curriculum can claim to be strictly learner-centred, unless consideration is given to students' perceptions and needs in the learning process. Kamgar and Jadidi (2016) stated

that learners' opinions, views, and preferences of the learning process are vital to a student's success or failure in learning a foreign language. Additionally, Ghanizadeh and Mirzaee (2012) argued that the more we understand about learners' individual methods and views, the more effective and fruitful our interventions would be; accordingly, language instructors and syllabus designers should pay attention to their learners' preferences and perceptions. This thesis aims to hear the voices of EFL learners regarding their perception and experience of self-assessment and its impact on their self-regulatory skills and critical thinking skills.

3.4. Conclusion

This chapter has reviewed literature on the role and impact of self-assessment on EFL learners' self-regulated learning and critical thinking. It began by discussing the developing concepts of self-regulated learning and critical thinking. It presented evidence suggesting that, through self-regulatory processes, learners can take charge of their own learning and thus play an active role in their learning, from setting goals to evaluating performance. However, a number of factors emerged that may influence the efficacy of that process.

This chapter has also provided evidence suggesting that, through recognising, adopting, and implementing cognitive skills and dispositions of critical thinking, students can enhance their learner autonomy and prepare to succeed in life. The evidence suggested that critical thinking and self-regulated learning had to be connected with higher-order and metacognitive strategies. Learners need to be actively involved in the learning process. The rationale behind activities should be explicit and linked to practical examples of how ideas might emerge in practice. Learners can thereby acquire a better understanding of their learning objectives. That improved understanding, in turn, can lead to the recognition of their own knowledge and abilities, and could promote the development of learning strategies through developing self-regulated learning and critical thinking.

Simplistic notions of self-assessment, self-regulated learning, and critical thinking linked to the use of strategies are unlikely to lead to the intended learning benefits. This chapter presented evidence of the significance of adopting a sociocultural perceptive instead of a constructivist perceptive regarding self-assessment. Social interaction and negotiation of meaning are key to achieving the intended learning benefits. Thus, feedback, from teachers, peers, or learners themselves, plays a vital role in achieving the anticipated

learning benefits; generating and managing feedback, in turn, enhances self-regulated learning skills and, thus, enhances learners' critical thinking.

The key message from this chapter is that perceptions, opinions, and preferences of learners about the learning process are always important to the success or failure of any student's attempts to acquire a foreign language. Language instructors and syllabus designers must therefore pay attention to their learners' preferences and perceptions. Accordingly, this thesis aims to hear the voices of EFL learners regarding their perception and experience of self-assessment and its impact on their self-regulatory skills and critical thinking skills. The following chapter discusses the methodology employed in this research.

Chapter 4 : Methodology of the Research

4.1. Introduction

As Chapter 3 reviewed the literature on assessment, self-regulated learning, and critical thinking, Chapter 4 discusses the used methodology to conduct this research. This chapter aims to explain and justify the selection of the research philosophy and research design. It aims to describe the methods used to collect and analyse the data in this research with the purpose of addressing the following research questions, as were described in Chapter One:

1) How do EFL students perceive the implementation of self-assessment in speaking classes, especially in relation to their speaking language achievement?
2) What influence does self-assessment in speaking have on EFL students regarding:
    a. their self-regulated learning in English language speaking?
    b. their critical thinking skills?
3) What is the association between students' perceptions regarding the influence of self-assessment in speaking on their self-regulated learning and their critical thinking?

This chapter also provides details about the institution where this research took place and outlines the procedures followed to obtain approval to collect the data from this institution and participants, according to the University of Glasgow's ethical standards along with the British Educational Research Association's (BERA) ethical standards for research involving human subjects. Additionally, this chapter explains the process of data collection following the university and offers some details and justifications for the sample selected.

4.2. Research approach

The philosophical stance plays a significant role in determining the research methodology (Creswell, 2009). Thus, identifying fundamental philosophical assumptions is essential to understanding and conducting a research project (Coe, 2017). The importance of identifying these philosophical assumptions lies particularly in designing and supporting the research framework (Hitchcock & Hughes, 2002). The research framework, in turn, guides researchers in the social sciences, affecting their choices and adoption of different methodological approaches (Gibson & Brown, 2009). Research philosophy includes a

range of paradigms (Lincoln & Guba, 2003). The following paradigms, as shown in Table 4-1, are commonly recognised in social research: positivism, critical realism, interpretivism, postmodernism, and pragmatism (Saunders et al., 2019, p. 144). A paradigm, according to Guba and Lincoln (1994), may be characterised as "a set of basic beliefs or assumptions [...] It represents a worldview that defines for its holder the nature of the world, the individuals in it and the range of possible relationships to that world and its part" (p. 108).

When considering how to approach this study, I began to seek the research paradigm that would be most appropriate, i.e. the paradigm that would allow me to fulfil the purpose of this research and address the research questions. The review of the literature (Chapter 3) revealed a research gap regarding the impact of self-assessment on self-regulated learning and critical thinking from the students' perspective in EFL speaking classes in Saudi Arabia. Therefore, this thesis aims to listen to the voices of EFL learners regarding their perceptions and experience of self-assessment and its impact on their own self-regulatory skills and critical thinking; this focus on learners' voices is appropriate, because learners' opinions, views, and preferences regarding the learning process are vital to the success or failure of efforts to learn a foreign language (Kamgar & Jadidi, 2016). In other words, in this research, I am interested in exploring and understanding EFL learners' perceptions of the implementation of self-assessment during speaking classes and how self-assessment affects their language performance, self-regulated learning, and critical thinking. What matters most in this research are the participants' own experiences and perspectives, which are unlikely to stay static and fixed and are what comprise the participants' reality; in short, the research is grounded on knowledge of an individual's lived experience. Accordingly, this research employs an interpretative paradigm to reach the study's goals.

Table 4-1 Summary of the five fundamental philosophical research paradigms in social science research (Saunders et al., 2019, p. 144)

Research Philosophy	Positivism	Critical Realism	Interpretivist	Postmodernism	Pragmatism
Ontology: the researcher's view of the nature of reality or being	Real, external, independent. One true reality (universalism) Granular (things) Ordered	Stratified/layered (the empirical, the actual and the real). External, independent Intransient. Objective structures. Causal mechanisms	Complex, rich. Socially constructed through culture and language. Multiple meanings, interpretations, realities. Flux of processes, experiences, practices	Nominal. Complex, rich. Socially constructed through power relations, Some meanings, interpretations, realities are dominated and silenced by others. Flux of processes, experiences, practices	Complex, rich, external 'Reality' is the practical consequences of ideas. Flux of processes, experiences and practices
Epistemology: the researcher's view regarding what constitutes acceptable knowledge	Scientific method. Observable and measurable facts. Law-like generalisations Numbers. Causal explanation and prediction as a contribution	Epistemological relativism. Knowledge historically situated and transient. Facts are social constructions. Historical causal explanation as a contribution	Theories and concepts too simplistic. Focus on narratives, stories, perceptions and interpretations. New understandings and worldviews as a contribution	What counts as 'truth' and 'knowledge' is decided by dominant ideologies. Focus on absences, silences and oppressed/ repressed meanings, interpretations and voices Exposure of power relations and challenge of dominant views as a contribution	The Practical meaning of knowledge in specific contexts. 'True' theories and knowledge are those that enable successful action. Focus on problems, practices and relevance. Problem solving and informed future practice as a contribution
Axiology: the researcher's view of the role of values in research	Value-free research. Researcher is detached, neutral and independent of what is researched. Researcher maintains an objective stance	Value-laden research. Researcher acknowledges bias by world views, cultural experience and upbringing. Researcher tries to minimise bias and errors Researcher is as objective as possible	Value-bound research. Researchers are part of what is researched, subjective. Researcher interpretations key to contribution. Researcher reflexive	Value-constituted research. Researcher and research embedded in power relations. Some research narratives are repressed and silenced at the expense of others Researcher radically reflexive	Value-driven research. Research initiated and sustained by researcher's doubts and beliefs. Researcher reflexive
Data collection techniques most often used	Typically deductive, highly structured, large samples, measurement, typically quantitative methods of analysis, but a range of data can be analysed	Retroductive, in-depth historically situated analysis of pre-existing structures and emerging agency. Range of methods and data types to fit the subject matter	Typically inductive. Small samples, in-depth investigations, qualitative methods of analysis, but a range of data can be interpreted	Typically deconstructive – reading texts and realities against themselves In-depth investigations of anomalies, silences and absences. Range of data types, typically qualitative methods of analysis	Following research problem and research question. Range of methods: mixed, multiple, qualitative, quantitative, action research. Emphasis on practical solutions and outcomes

Within the interpretivist paradigm, reality is regarded as multiple and complex, with various and distinct interpretations for a single phenomenon. Interpretivists emphasise the significance of an individual's subjective experience rather than viewing social reality as a single fixed reality. In contrast, positivists believe that there is one singular truth, one reality; thus, positivists presume that reality could be generalised and is applicable to other situations (Croker, 2009). Accordingly, positivists tend to adopt quantitative approaches and mostly use instruments that allow them to objectively collect data from large samples to measure the research issue (Picciano, 2004). However, although my research includes an intervention with students and the use of quantitative data, I did not follow the positivist paradigm, which would usually be the choice of a researcher who follows objective assumptions rather than subjective. I was more interested in the participants' subjective experiences and their perceptions of the intervention and its impact than in trying to measure the intervention's results.

The interpretive paradigm views knowledge as "personal, subjective and unique" (Cohen et al., 2007, p. 7) and is defined by its emphasis on the individual. The interpretive paradigm, in other words, focuses on the explanation and comprehension of the unique individual case rather than on the general; it is interested in comprehending how the individual generates, alters, and explores the world that the individual inhabits. Interpretivists tend to adopt qualitative research, which is generally defined as any type of research that generates results that are not quantifiable through statistical techniques or other quantification methods (Davis, 1995). This preference is due to interpretivists establishing the foundation for data gathering as the collection of rich data that can support investigating, characterising, and promoting the understanding of organisational social and psychological processes and the social contexts in which individuals find themselves (Berg, 2004; Bogdan & Biklen, 1997). Interpretivism imposes on researchers' constraints regarding the use of the ways of the natural scientist and how they interact with their participants during deep investigation of the research issue.

In light of the above, this research employs an interpretative paradigm to represent my viewpoint and the study's goals. To illustrate, I view learning as the result of a dynamic interaction between individuals, not as an objective incident unrelated to the learners or their environment. I believe that the context in which a person lives influences their learning process and perspectives. Additionally, participants' views of the problem may differ according to their own perception and experience, and those differences may ultimately result in a different interpretation. For all these reasons, this research employs

an interpretive paradigm; in particular, a phenomenological approach: a form of interpretive paradigm focused on comprehending individuals' lived experience (Cal & Tehmarn, 2016).

The methodologies of the interpretivist paradigm usually illustrate actions, through qualitative research methods, that the participants exhibit or explain through open-ended interviews, classroom observations, and focus groups (Bryman, 2016; Scotland, 2012). Qualitative research is often concerned with how individuals perceive, comprehend, and make sense of their surroundings. By examining the views of individuals who have personally experienced a certain occurrence, a researcher may acquire and comprehend different perspectives on the phenomena. Kumar (2019) emphasised the value of qualitative research in providing a thorough understanding of individuals' experiences.

However, social scientists have acknowledged that quantitative and qualitative research methods could be combined, merged, or integrated into a single research for rational purposes (Creswell, 2010). This argument highlights the fact that, despite the fact that various research methodologies have typically been associated with certain paradigms, this connection between research paradigms and research methods is neither essential nor sacred (Johnson & Onwuegbuzie, 2004). Crotty (1998) argued that, if a method suits their purposes, researchers could adopt any theoretical perspectives and utilise any methods of research. Similarly, Johnson and Onwuegbuzie (2004) argued that significant disagreements in epistemological views should not hinder a qualitative researcher from using data collection methods that are often linked with quantitative research, and vice versa. Building on the two previous opinions, McChesney and Aldridge (2019) argued for including both qualitative and quantitative methods in mixed methods research in accordance with a single and particular research paradigm. This position allows for a more diverse and purposeful selection and integration of paradigms and/or techniques in order to achieve a study's objectives (McChesney & Aldridge, 2019).

In this research, the context in which the study was conducted, the Kingdom of Saudi Arabia (KSA), and the aim that the study pursued, which was to explore and measure learners' perception of their critical thinking and self-regulated learning, both had a substantial impact on the method choice. To illustrate, in the KSA, quantitative methods, particularly standardised tests and questionnaires, are two of the most important tools for assessing education and educational institutions; these tools collect numerical data that can be converted into useful statistics for evaluating students' achievement (ETEC, 2018). Quantitative methods thus contribute to identifying the most critical and effective ways to

improve education (ETEC, 2018). In addition, in the KSA, standardised tests and questionnaires are also important for exploring attitudes, views, and behaviours (ETEC, 2018). In addition to the impact of the context, the literature review indicated that perceptions of self-regulated learning can be collected and measured through questionnaires, and that critical thinking can be measured through standardised testing (see Chapter 3 Section 3.2). However, focusing on the numerical data alone would not have been suitable for this research and would be unlikely to support the complexity of the research issue; that is, EFL learners' perceptions towards self-assessment in speaking classes and the impact of those perceptions on their speaking skills, self-regulatory skills, and critical thinking. This research issue, in short, is mainly concerned with perceptions and attitudes, which are difficult to measure, but can be explored through qualitative methods (Creswell, 2012). Accordingly, this research adopted a mixed methods research design under the interpretivist paradigm instead of solely employing quantitative methods to explore students' perceptions.

According to Creswell (2012) and Bryman (2014), combining quantitative and qualitative data collection methods results in a better understanding of the research questions than using either method alone. Thus, the first rationale for using mixed methods is triangulation. Triangulation can be defined as the use of several methods during the investigation of a single phenomenon (Denzin, 1978). According to Denzin (1978), triangulation in research may be classified as follows: data triangulation, investigator triangulation, theory triangulation, and methodological triangulation. Data triangulation refers to the utilisation of several sources of evidence which are brought together to build a more nuanced understanding of a research topic. Investigator triangulation refers to the practice of using several researchers in the context of a study. Theory triangulation refers to the process of interpreting data using several theories (Denzin, 1978). Finally, methodological triangulation refers to the investigation of a phenomena via the use of various techniques (Denzin, 1978). Methodological triangulation was used in the present study. Additionally, this study adopted a particular type of methodological triangulation. Two types exist: within-methods triangulation, which involves using a combination of qualitative or quantitative methods, and methods triangulation, which involves using both quantitative and qualitative methods. The use of methods triangulation in this research is discussed later in Section 4.6.3. In this study, I used a combination of quantitative and qualitative data; according to Denzin (1978), integrating data in this way can decrease the bias of any individual methods.

The purpose of this research is not to develop statistical generalisations about other learners. Nevertheless, it does aim to develop case-to-case translation or transferability, based on Firestone (1993) three models of generalisability, through providing a comprehensive examination of participants' assessment experiences and perceptions of the impact of self-assessment on their self-regulatory skills and critical thinking within a particular context. The issue of transferability is discussed later in Section 4.6.4.

This indicates another rationale for the use of mixed methods in that the combination of quantitative and qualitative methods extends, explains, illustrates, and complements the results of each individual methods, thus avoiding the disadvantages of utilising only a single method by itself (Bryman, 2014; Creswell, 2009; Creswell, 2012; Johnson & Onwuegbuzie, 2004). Additionally, mixed methods research assists a researcher to gain more insightful and balanced findings (Creswell, 2009; Creswell, 2012; Fraenkel et al., 2012). Punch (2009) provided an illustration of how both approaches may boost research:

> Quantitative research brings the strengths of conceptualizing variables, profiling dimensions, tracing trends and relationships, formalizing comparisons and using large and perhaps representative samples. On the other hand, qualitative research brings the strengths of sensitivity to meaning and to context, local groundedness, the in-depth study of smaller samples, and greater methodological flexibility which enhances the ability to study process and change. (p. 290)

This research adopts a certain type of mixed methods research: the embedded design. The purpose of using an embedded design, according to Creswell (2011), is to gather quantitative and qualitative data sequentially and to have one form of data that acts as a support for another. Creswell (2011) explained that the process of an embedded study is to gather quantitative and qualitative data in the course of a single research project, with both datasets being examined independently and addressing different research objectives. For instance, in the case of the present research, whether the intervention had an impact on the outcomes is assessed using the quantitative data, while the perceptions of participants and attitudes about the intervention are assessed using the qualitative data (Creswell, 2011).

This research employs an interpretivist paradigm and is concerned with investigating the effect of self-assessment in speaking classes on EFL learners' perceptions of their self-regulatory skills, critical thinking, and speaking performance in the KSA. The research site and English programme are discussed in the next section.

4.3. Research site

The study was conducted in a single Saudi higher education institution (HEI), a national university in the western region of the KSA. This HEI offers an EFL programme for preparatory year students that aims to build students' English skills in key areas both academically and personally. The English Language Institute (ELI) at this HEI delivers two levels of intensive English language courses (ELI_101 and ELI_102). These courses aim to improve learners' English proficiency in four key areas (listening, reading, writing, and speaking); to reinforce these abilities with linguistic and lexical abilities; and to develop thinking skills, presentation skills, and sub-skills. The curriculum design process is term-based, with each term consisting of a total of 14 instructional weeks per level. Like all other courses in the preparatory year, successful completion of the course requirements enables students to be admitted into undergraduate programmes. It is noteworthy that specialists such as Derwing et al. (2004) and Zhang (2009) indicated that traditional language teaching approaches in which the four abilities are taught individually are unsatisfactory, and argued that transition to a more integrated and communicative approach of learning is essential.

The research was carried out in level ELI_102 in the second term of the academic year. As described in the institution's policy documents, ELI_102 is an intensive English course which aims to develop language proficiency beginning at a B2 level and moving into the C1 level on the Common European Framework of Reference for Languages (CEFR), which is within the band of being an independent/proficient user (Canadian benchmarks CLB 8-9-10) (ELI-Curriculum-manual, 2018/2019). This progress from a B2 level into the C1 level marks a transition from being an independent language user to being a proficient language user (see Appendix A for descriptors of the common European framework of reference for language levels). This course provides students with the opportunity to develop their language abilities and thinking and presentation skills. Students in this course apply their language skills in a variety of activities, including reading, writing, speaking, and oral presentations incorporating a variety of formats such as reading passages, articles, videos, lectures, and audio materials.

In the course, instead of a textbook, the students are provided with a course kit that is considered to be a learning and practice guide that should be supported with many extra-curricular resources and tasks. This course kit consists of 14 pacing guides, each for a week of the 14 instructional weeks in the semester, each with a different theme and

different main and skill-specific student learning objectives (SLOs). Each day of the week is designated to the development of a certain skill (grammar & vocabulary, listening, reading, writing, and speaking) (see Appendix B for a sample of the pacing guides). Each instructional week begins with class of grammar & vocabulary related to the week's theme and ends with a speaking class.

Moreover, the classroom activities in the programme include discussions, vocabulary building, writing and speaking exercises, pronunciation and grammar lessons. Through this course, students learn how to make visual outlines and information graphs, how to summarise, analyse, and brainstorm, and how to use a variety of thinking skills in different contexts. The pedagogy evident in the institution from learners' perceptions is discussed later in Chapter Seven.

Furthermore, for the measurement of the learners speaking performance the teachers use the institution's specific rating criteria and assessment scale that include descriptions of eight components of a speaking performance including: fluency, pronunciation, topic management, content, grammar (accuracy), vocabulary, communicative effectiveness, and organisation & delivery.

The target population of this study and sampling methods are discussed in the next section.

4.4. The target population and sampling methods

Making an informed choice about participant selection involves taking several critical factors into consideration. Cohen et al. (2007) suggested considering the following factors in making decisions about sampling: the population sample's size, the sample's accessibility, and the sampling strategy. This research was conducted at a Saudi HEI. Similar to other universities in the KSA, the institution's foundation year includes rigorous English language programmes. As discussed in the previous sections of this chapter, this was a mixed methods research with both quantitative and qualitative parts. Therefore, different sampling strategies were adopted to determine the sample for each part.

The intervention and quantitative parts of this research adopted the sampling strategy of random selection. Proper sampling is crucial in quantitative sampling, since quantitative research relies on drawing a representative sample of the target population so that the findings may then be generalised to the population (Shah & Al-Bargi, 2013). Nonetheless,

it is important to note that this research does not intend to make statistical generalisations. In random selection of participants, a researcher draws a sample from a population at random (Creswell, 2012; Herbert et al., 1989). One section of 35 students was randomly selected from all of the ELI_102 students registered for the year of 2019. The recruitment of participants is discussed later in Section 4.6.3.

With regards to the sample for the qualitative part of this research, which included self-assessment proformas, audio recordings, and interviews, the adopted sampling strategy is referred to as criterion sampling. In qualitative research, the sample size should not be so small as to preclude data saturation, defined as the moment at which no new themes or codes originate from data (Braun & Clarke, 2021), but it should not be so large that it is impossible to conduct a thorough, case-based analysis (Bryman, 2014). Accordingly, this research employed non-sequential approaches to sampling that are referred to as fixed sampling strategies. With fixed sampling strategies, the sample is set early on in the study, and there is no sample increase as the research progresses (Bryman, 2014). In particular, this study follows criterion sampling, a type of fixed sampling strategy identified by Patton (1990) and Palys (2008), in which all individuals who meet a certain criterion are included. In this case, the criterion was that participants must be students who participated in the intervention from the beginning, since any student who enters the sample during the intervention would have a different experience than students who were included from the start.

The following section addresses the process of recruiting the participants and the ethical considerations in this research.

4.5. Ethical considerations

Ethical issues are directly connected to the integrity of a research project and cannot be disregarded (Bryman, 2016). Ethics in research consist of the norms of behaviour that guide researchers' actions with regard to how they preserve the rights of study participants or any individuals who may be impacted by the research (Guillemin & Gillam, 2004). The researcher developed this thesis in accordance with the University of Glasgow's Ethics Committee's standards (UofG, 2018), which necessitate the submission of an ethical approval application. In line with the university’s requirements, the researcher filled in an application to the University of Glasgow's College of Social Sciences Ethics Committee that highlighted the ethical concerns to be considered and how they would be handled

before any fieldwork or data gathering began. The application for ethical approval was approved prior to data collection (see Appendix C). In the following paragraphs, I provide examples of how I have adhered to the University of Glasgow's ethical standards, as well as to the British Educational Research Association's (BERA) ethical standards for research involving human subjects. Permission for conducting the empirical study was also acquired in advance from the university and from all participants in this research.

Prior to conducting any research or beginning data collection, gaining access to an investigation or formal setting, such as the classroom, is critical and can be done through obtaining permission from "gatekeepers" (Creswell, 2009). In this case, the gatekeeper was an educational establishment. Creswell (2009) also suggested that it is better to create a brief proposition for review by management in regard to the research. This proposal should contain the study's goals, significance, and purpose for the administration, and include the name of the dean of the institution to which the researcher is seeking access (Cohen et al., 2007; Creswell, 2009). Accordingly, I contacted the dean and vice dean of the ELI at this Saudi HEI in order to attain their permission to conduct the study.

Moreover, naturalistic researchers, such as Lofland and Lofland (1985) in Hoepfl (1997), have proposed and emphasised the significance of asking individuals to provide access to their lives, experiences, and perceptions based on justifications for the study's objectives. Thus, prior to the implementation of the intervention and data collection, the prospective participants, who were EFL students in this case, were informed of the study's nature, purpose, and significance. An introductory PowerPoint presentation was given to the whole class, and then the participant information sheet (PIS) and consent form (see Appendix D and E) were presented. Both these documents were in the students' first language (Arabic). This content clarified and explained what was meant by self-assessment and outlined the study's primary purposes and benefits in addition to some key concepts related to the study (e.g. critical thinking and self-regulated learning). The content also offered a brief outline of the purpose of the study, its instruments, and the intervention. Then, the participant information sheet and consent form were distributed to the students and collected the following day so that the students could read both documents thoroughly and make the decision whether to participate in the comfort of their own homes.

Each student was informed that their participation was optional and that they could withdraw at any time. Each student signed the consent form and acknowledged that the first and last speaking activities and interview would be recorded. Students were also informed that, as participants, they would be identified only by an ID number that would

then be de-identified and replaced by a code in any report or publication. Students were also informed that anyone who did not want to participate would not be affected by the study or the intervention, as any time spent assessing themselves would be done during free time for the students, which they regularly have at the end of each class anyway; during that time, the students are free to do whatever they want to do as long as they are speaking in English. The data collection instruments and process, and the participants' recruiting process are discussed in the next section.

4.6. The data collection process and its reliability and validity

This research utilises primary, or original, data as well as extant literature. The materials reviewed for this thesis include books, journal articles, and online resources; these sources provided the study's contextual and theoretical framework. Primary data are data that have been collected specifically for the purpose of the research (Neuman & Kreuger, 2003). In the current research, the primary data were collected from the ELI at the aforementioned Saudi HEI via pre-post tests, self-assessment proformas, audio recordings, and semi-structured interviews.

4.6.1. The intervention

The present study aimed to explore EFL learners' perceptions of self-assessment in speaking classes and its impact on their critical thinking, self-regulated learning, language performance, and achievements. The study was conducted with EFL students in ELI_102 (see Section 4.3) during the second term of the academic year 2018/2019 during speaking classes. These speaking classes were held and taught as usual by the teacher with little reference to the intervention. As mentioned in Section 4.5, prior to the implementation of the intervention, an introductory PowerPoint presentation informed the prospective participants of the study's nature, purpose, significance, instruments, and intervention, and gave them the opportunity to enquire about and discuss the intervention. It also illustrated the nature of self-assessment as well as key concepts related to the study. The students participated in individual, paired, or group speaking activities and then spoke or presented in front of the class. The intervention took place during speaking classes in a total of eight speaking classes. Each class was approximately three hours long. The classes were included in weeks 3, 4, 6, 7, 9, 10, 11, and 13 of ELI_102. As mentioned in Section 4.3, each week covers a new theme and pacing guide which include new objectives, grammar lessons, and word lists and a different speaking task (see Table 4-2 Speaking tasks

administered during weeks 1 to 8 of ELI_102). During the intervention, I attended the speaking classes with the learners and was available to support them as well as discuss and respond to their concerns regarding the intervention when needed.

By the end of the intervention, all the students participating in the study had engaged in a total of eight speaking classes and completed all eight speaking activities.

Table 4-2 Speaking tasks administered during weeks 1 to 8 of ELI_102.

Weeks	Date	Theme of the Week	Speaking Activity
Week1 (***Week3)***	17-2-2019	e-Business revolution	Prepare a short presentation based on a listening task. Students were required to listen to an audio recording on e-business, take notes, and summarise the content for their classmates.
Week2 (***Week4)***	24-2-2019	Online shopping	Narrate a personal experience with online shopping. Students were required to tell their classmate about their own experience with online shopping.
Week3 (***Week6***)	10-3-2019	Taking a selfie	Talk about taking a selfie. Students were required to talk about their best or worst selfie, or give tips on how to take the perfect selfie.
Week4 (***Week7***)	17-3-2019	Second language learning	Interview a classmate on second language learning. In pairs, students were required to take turns interviewing each other about their experience learning a second language.
Week5 (***Week8***)	24-3-2019	International tests of English	Discuss international tests of English. Students were required to talk about and discuss their own experiences or information regarding international tests of English and to compare the tests.
Week6 (***Week9***)	7-4-2019	Best business leaders ever	Discuss the best business leaders ever. Students were required to discuss, from their perspective, the business leaders who were the best in the world.
Week7 (***Week10***)	14-4-2019	Multiple intelligences	Prepare a presentation based on listening task. In groups, students were required to listen to an audio recording on multiple intelligences, take notes, then summarise the notes for their classmates.
Week8 (***Week11***)	21-4-2019	Saudi Vision 2030	Prepare a presentation on Saudi Vision 2030. Students were required to give a presentation on Saudi Vision 2030 with the use of a mind map that included the main points.

Manso-Vázquez et al. (2018) self-regulated learning (SRL) model was used to design and develop the intervention as a means to optimise EFL learners' SRL through self-assessment during the speaking classes. As referred to in Chapter 3, this model is an integrated model that combines the three most often used self-regulated learning models:

Zimmerman (2002) model, Pintrich (2004) general model, and Winne and Hadwin (1998) information processing model. This model includes three phases: (1) forethought, planning, and activation; (2) performance, monitoring, and control; and (3) evaluation, reflection, and reaction. Each phase targets five different areas: cognition, metacognition, motivation/affect, behaviour, and (social and environmental) context.

The first phase (forethought, planning, and activation) includes the processes that occur when learners prepare to learn or perform (see Figure 4-1). In the intervention, during this phase, the teacher presented the speaking task to the learners and discussed its requirements with them. The learners were advised to use a self-assessment proforma to plan and set their goals. This phase included the application of different learning strategies, as the teacher defined and described the actions that the learners would use to perform in planning and preparing for the task. The phase addressed all five target areas as follows: selecting sources, taking notes, and summarising (cognition area); defining the goal and task (metacognition area); discussing the difficulty of the task and level of interest with the teacher and classmates (motivation/affect); estimating the time required (behaviour); and discussing the task (context).

During the second phase (performance, monitoring, and control), the learners were asked to use self-assessment proforma to monitor their work against the speaking criteria: fluency, pronunciation, topic management, content, grammar (accuracy), vocabulary, communicative effectiveness, and organisation & delivery that was included in the self-assessment proforma and was also used by the teacher to assess their performance. The learners were also asked to audio-record themselves while they presented or talked in class, so they would be able to listen to their own performance afterward and self-assess themselves with the use of the proforma. Throughout the second phase, learners applied different strategies to enhance their learning and attention and concentrate on task achievement within the five target areas (see Figure 4-1) as follows: practising vocabulary and grammar control, memorising through imagery, and summarising meaning (cognition area); self-monitoring through the self-assessment proforma (metacognition area); using the checklist on the self-assessment proforma (motivation / affect); exercising time control (behaviour); and collaborating and cooperating for task achievement (context).

During the third and last phase (evaluation, reflection, and reaction), learners used different strategies to respond to critical issues related to their level of performance, goal achievement, and task type. The five target areas were addressed as follows (see Figure 4-

1): brainstorming ideas for enhancing performance through the self-assessment proforma (cognition); evaluating performance through the self-assessment proforma (metacognition); reacting after evaluation of performance and making positive statements through the self-assessment proforma (motivation/affect); evaluating the allocated time and recognising strengths and weaknesses (behaviour); and discussing and evaluating the task (context).

Figure 4-1 Integrated self-regulated learning model (adapted from Manso-Vázquez et al. (2018)

During the intervention, the students were given the freedom to decide whether to discuss their performance and assessment with their classmates or teacher; this option was given to avoid some sociocultural challenges specific to the Saudi culture. These challenges include students' fear of negative comments (Al-Haqwi et al., 2012) and the relationship between teachers and students in Saudi culture, wherein mastery of knowledge is reserved for the instructor (Al-Wassia et al., 2015). In short, Saudi students tend to be fearful of engaging in debates with their teachers, and teachers tend to dislike and may even resent questioning and discussion from the students (Kariri et al., 2018). Alternative forms of assessment, like

peer assessment, may therefore be less appreciated by teachers and students in such cultures (Yan & Cheng, 2015).

4.6.2. Research instruments

As noted before, both quantitative and qualitative data collection instruments were selected and used to collect data in this study in order answer the research questions and provide thorough knowledge of the phenomenon being studied. For the quantitative data in this mixed methods research, two instruments were utilised: pre- and post-tests, and students' audio recording of their speaking activities. The quantitative data collected via these instruments helped to measure the impact of self-assessment on students' perceptions regarding their critical thinking and self-regulation before and after the intervention. As for the qualitative data in this mixed methods research, the instruments used were a self-assessment proforma, and semi-structured interviews. The qualitative data were intended to provide a better understanding of students' perception regarding the impact of self-assessment on their thinking, self-regulation, and speaking. All the adopted instruments in this study (the pre- and post-test, self-assessment proforma, audio recording, and semi-structured interviews) are outlined in the following subsections in the order in which they were used.

4.6.2.1. Pre- and post-tests

According to Creswell and Creswell (2017), a pre-test is a test that assesses some traits or features of participants prior to their participation in an intervention. Following the intervention, the post-test then reassesses those traits or features of participants; in this research, the post-test aimed to assess learners' critical thinking, self-regulated learning, and speaking achievement at the end of the semester after the intervention. The two instruments utilised were specifically pre-post surveys, which collected the quantitative data. The quantitative data, in turn, helped identify and explore the impact of self-assessment on students' perceptions regarding their self-regulated learning through the Self-Regulation of Learning Self-Report Scale (SRL-SRS), and critical thinking through the Watson-Glaser II Critical Thinking Appraisal (CTA) before and after the intervention; in other words, at the beginning and the end of the academic semester.

4.6.2.1.1. The Self-Regulation of Learning Self-Report Scale (SRL-SRS)

This research aims to explore students' perceptions of their self-regulated learning before and after the intervention. Thus, the Self-Regulation of Learning Self-Report Scale (SRL-SRS), which was translated into the students' first language of Arabic (permission for translation was granted from the authors) (see Appendix F), was the first instrument utilised in the research. Self-report questionnaires are the most often used technique for measuring SRL, as noted by Perry and Winne (2006). Self-report questionnaires offer important information on how learners perceive their ability to self-regulate their learning.

The SRL-SRS has been used in several contexts, including EFL contexts, to measure students' SRL (Saks, 2016; Zarei et al., 2016). It contains 50 items that were composed by Toering et al. (2012). The scale seeks to measure students' perception of self-regulated learning processes through measuring metacognition and motivation dimensions before and after the intervention. The scale results thus help to answer the second research question about what influence self-assessment in speaking has on EFL students' perceptions regarding their self-regulated learning in English language speaking classes. The metacognition dimension contains four sub-scales: planning, self-monitoring, evaluation, and reflection. The motivation dimensions consist of effort and self-efficacy. The subscales of planning and effort were based on the self-regulatory inventory by Hong and O'Neil Jr. (2001), and the self-monitoring subscale was adopted from the Self-Regulation Trait Questionnaire by Herl et al. (1999). Self-efficacy was assessed with items based on the Generalized Self-efficacy Scale (Hong & O'Neil Jr., 2001; Schwarzer & Jerusalem, 1995). The evaluation items were adopted from the evaluation subscale of the Inventory of Metacognitive Self-Regulation (Howard et al., 2000), and the reflection subscale was based on the reflection subscale of the Reflective Learning Continuum (Peltier et al., 2006).

The SRL-SRS consists of two sections. Each section targets one of the two dimensions with a total of 50 statements:

- First section: Targets metacognition dimensions and includes four subsections with a total of 30 statements.
    - *First subsection (Planning):* Aims to measure the degree to which an individual must plan their approach to a task in advance of their actions (includes 9 statements).

  - *Second subsection (Self-monitoring):* Aims to measure the degree to which an individual requires a self-checking system while doing a task in order to monitor achievement of goals (includes 8 statements).
  - *Third subsection (Evaluation):* Aims to measure the degree to which an individual must assess the process and results of their plan after execution (includes 8 statements).
  - *Fourth subsection (Reflection):* Aims to measure the degree to which an individual must reflect on their learning process, implying that they apply their information and expand their repertoire of methods, thus increasing their performance possibilities (includes 5 statements).
- Second section: Targets motivation dimensions and includes two subsections with a total of 20 statements.
  - *First subsection (Effort):* Aims to measure the degree of energy an individual exerts in working on a task (includes 10 statements).
  - *Second subsection (Self-efficacy):* Aims to measure the degree of individual confidence in their own ability to complete a task (includes 10 statements).

Each of the six subsections were measured using Likert scales of different points. A four-point Likert scale was used to measure four out of the six the subsections (planning, self-monitoring, effort, and self-efficacy). The learners marked their perception within four categories; the higher the number chosen, the more often they believed they performed a certain skill or ability (1 = almost never, 2 = sometimes, 3 = often, and 4 = almost always). The remaining two subsections were measured on a five-point Likert scale, though in each of these two sub-sections the learners marked their perception within two different categories. In the evaluation sub-section, the participants marked their perception on a 5-point scale in which the higher the number chosen, the more often they believe they performed a certain skill or ability (1 = never, 2 = rarely, 3 = sometimes, 4 = often, and 5 = always). In the reflection subsection, the participants marked their perception on a 5-point scale in which the higher the number chosen, the more they agreed with the given statement (1 = Strongly Disagree (SD), 2 = Disagree (D), 3 = Neither (N), 4 = Agree (A), and 5 = Strongly Agree (SA)). Both the validity and reliability of the SRL-SRS have been proven in different studies and are discussed later in Section 4.6.3.

4.6.2.1.2. Watson-Glaser II Critical Thinking Appraisal (CTA)

As referred to in Chapter 3, the California Critical Thinking Skills Test and the Watson-Glaser Critical Thinking Appraisal are the two commonly used tests to measure critical thinking abilities in the field of language learning. Both tests are designed for graduate and undergraduate university students. However, the Watson-Glaser Critical Thinking Appraisal test was used in the current study to measure the students' critical thinking abilities since the California Critical Thinking Skills Test is analysed by Insight Assessment, the providers of the test; in the case of the present study, that would have compromised the confidentiality of the participant information and results. The Watson-Glaser II Critical Thinking Appraisal was initially developed by Goodwin Watson and Edward Glaser in 1925. It is a multiple-choice assessment instrument that has been developed specifically to evaluate chosen key critical thinking abilities. It has been used in academic contexts to assess the extent to which students have improved their critical thinking abilities as a consequence of coursework or instructional programmes. It has also been used as a selection and developmental tool in organisations. The test has been applied specifically in several studies in the EFL context to measure students' critical thinking (Bagheri & Ghanizadeh, 2016; Ebrahimi & Moafian, 2012; Ghanizadeh & Mirzaee, 2012; Jafari et al., 2015). The test has 40 multiple-choice items in 5 sub-test areas that can be completed within an estimated 35 to 50 minutes. It aims to measure CT through measuring the ability to logically analyse assumptions, evaluate the strength of an argument, and draw logical conclusions from the information provided. The test results thus help to answer the second research question, about what influence self-assessment in speaking has on EFL students' perceptions regarding their critical thinking in English language speaking classes.

It consists of five sections, and each section targets a single critical thinking skill with a total of 40 statements:

- First section *(Infer)*: Aims to measure the skill of inference through asking learners to identify the degrees of truth or falsehood of an inference with 5 statements.
- Second section *(Recognise Assumptions):* Aims to measure the skill of acknowledging underlying assumptions or presuppositions with 12 statements.
- Third section *(Deduce):* Aims to measure skill of deduction through asking learners to decide whether or not certain findings necessarily follow with 5 statements.
- Fourth section *(Interpret)*: Aims to measure skill of interpretation through asking learners to evaluate evidence and determine whether generalisations or conclusions are appropriate with 6 statements.

- Fifth section *(Evaluation of Arguments):* Aims to measure skill of evaluation of arguments by determining whether learners can distinguish between strong and relevant arguments and arguments that are weak or irrelevant with 12 statements.

The paper and pencil version of the test was administered in English. According to the Watson-Glaser II test manual, "directions and items were written at or below the 9th grade reading level"; that is, at the B2 reading level or below (Watson & Glaser, 2010). This level is at or below the expected reading level of undergraduate students in ELI_102 participating in the study (expected level is B2/C1 independent/proficient user) (ELI-Curriculum-manual, 2018/2019). Thus, participants were not expected to experience any difficulties with the test instructions and items. In addition, the test manual states that the Watson-Glaser II should be given in the examinee's first language if feasible; however, an Arabic version is not available for the test, and translating the test is not permitted. Regarding the administration of the test, though no special training is required to administer the Watson-Glaser II, the administrator must be capable of conducting standard examination procedures and fully familiarised with the administration instructions and test materials prior to administering the test in order to guarantee accurate and reliable results (Watson & Glaser, 2010). Students' participation and test results had no effect on their grades in the course. Both the validity and reliability of the Watson-Glaser II Critical Thinking Appraisal have been proven in several studies and are discussed later in Section 4.6.3.

4.6.2.2. Audio-recordings

As discussed in Chapter 3, listening to recordings of themselves can offer students a valuable opportunity to gauge their own linguistic competence. When audio recording is used, students can evaluate their own paralinguistic skills, including the tone and pitch of their voice, as an adjunct to their language parallelism and linguistic performance. Repeated use of this technique can allow learners to acquire a precise sense of their progress and performance, which is beneficial in terms of increasing learner confidence levels pertaining to oral productive skills. Thus, the intervention in the study includes audio-recording students while presenting, talking, or debating and using those recordings afterward to self-assess per predefined criteria included in the self-assessment proforma, as discussed in the following section. Each student was engaged in a total of eight speaking classes and activities, had 24 hours of speaking practice, and produced a total of eight audio-recordings. These recordings were utilised as the third instrument in the research to

measure student speaking ability and achievement at the beginning and the end of the intervention. The first and last (i.e. eighth) recordings were collected from all 10 students comprising the sample in the study.

4.6.2.3. Self-assessment proforma

Depending on the purpose of a study, self-assessment may take a variety of forms. Examples of forms of self-assessment include a questionnaire assessing general speaking skills; a reflective conversation; a learning journal for metacognitive reflection; or a classroom activity, such as a self-assessment proforma, where students utilise the same assessment standards or rating criteria as their instructors (Bachman & Palmer, 1989; Chen, 2008; Rivers, 2001). This research asked the students to use a unique criteria-referenced self-assessment proforma with a five-point Likert scale during their speaking classes, during speaking activities and after listening to recordings of each speaking across a total of eight speaking classes and activities (see Appendix H). Andrade and Cizek (2010) and Panadero and Alonso-Tapia (2013) argued that in order to conduct an adequate self-assessment of their work, students need to understand the criteria provided by their teachers. Thus, the creation of the proforma, which includes eight criteria, was driven by the institution's policy and approach of teaching and follows the same rating criteria and the scoring system and rubric used by teachers, the pacing guides, the week-specific SLOs, and the activity instructions. These eight criteria are fluency, pronunciation, topic management, content, grammar (accuracy), vocabulary, communicative effectiveness, and organisation & delivery.

Self-assessment proformas were used for "Tracing", which refers to a method introduced by Perry and Winne (2006) for gathering data by involving learners in tasks that provide information about learners' engagement with the tasks. Traces, or evidence gathered over time, may be used to provide precise, time-stamped descriptions of observed engagement between students and content (Perry & Winne, 2006). Thus, the self-assessment proforma can provide traces of what learners really did and what SRL strategies they actually used, rather than what learners claim to have done and used (Perry, 1998; Winne & Jamieson-Noel, 2002).

4.6.2.4. Interviews (semi-structured interviews)

Interviews are a valuable research tool for eliciting information about a phenomenon from a variety of respondents in order to accomplish research goals. Surveys can provide rich and valuable information (Cohen et al., 2018). Surveys are, however, guided by their nature and are fairly prescriptive; surveys do not provide responders with any freedom or the opportunity to seek explanation or engage in discussion (Cohen et al., 2018). Interviews, in comparison, may be beneficial and offer valuable insights on this front. According to Kvale (1996), interviews can be defined as the exchange of viewpoints between two individuals talking about a shared interest. The interviews in this research allowed me to establish a relationship of mutual trust with my interviewees and to acquire data that the other types of instruments would not have allowed me access to. The interviews also allowed the participants to share opinions and experiences and allowed me the opportunity to conduct a precise analysis of their answers (Wimmer & Dominick, 2013). Furthermore, interviews provide the researcher with a chance to obtain opinions, responses, or clarifications about certain actions on the part of the interviewees. As a result, this technique was critical to the study since it enabled a more in-depth investigation, formation of a more complete picture of the learners' perceptions, and a more comprehensive description of the study objectives. Additionally, Yin (2009) stated that interviews should be considered vital for any study with a small sample size as interviews add depth and deepen understanding regarding the issue under investigation. Again, that was the case for this study. The interviews in this study were semi-structured. I pre-planned questions based on the literature and the research question, but allowed the students to respond based on their own experiences and perspectives; while doing so, I provided opportunities for the exploration of topics significant to each learner.

The semi-structured interview format with students was utilised as the last data collection instrument in the study to collect the qualitative data. Semi-structured interviews are one of the interview types that Burns (2009) highlighted in educational research, the others being structured and unstructured interviews. In structured interviews, the researcher asks all participants the same set of control questions. In unstructured interviews, in contrast, the researcher makes use of spontaneous, open-ended questions. Combining these two types of interviews, semi-structured interviews start with a list of less strictly controlled questions that the researcher may supplement, omit, or alter based on the interview. Moreover, in a semi-structured interview, as explained by Bailey (2007):

> The interviewer uses an interview guide with specific questions that are organised by topics but are not necessarily asked in a specific order. The flow of the interview, rather than order in a guide, determines when and how a question is asked. (p. 100)

The semi-structured interview format was selected since it enables the researcher to have complete control over the interview while allowing for more freedom in terms of question sequence, probing, and in-depth examination of answers. This created the opportunity for the participants to engage in fruitful discussion about a certain topic and contributed in outlining the areas to be investigated. This decision thus enabled me to include questions pertaining to particular themes that I anticipated would arise, but might not naturally emerge during the interview (Gillham, 2000; Kvale & Brinkmann, 2009). Additionally, this structure enabled me to pose a question sooner, if necessary, even if the question was not on the interview questions order. Additionally, semi-structured interviews enabled me to ask follow-up questions and prompts throughout the interviews in order to elicit additional information from participants. The interviews consisted of 14 questions related to four themes that were predetermined; however, students had the opportunity to clarify, explain, and depart somewhat from the questions. The questions were prepared as open-ended questions. Prompts and sub-questions were then included to elicit more insights into the students' responses.

The key objective of the interview questions was to explore and discuss students' attitudes towards the intervention of self-assessment with the use of audio recording and self-assessment proforma in their speaking classes and how the self-assessment affected their performance, critical thinking, and self-regulation. The questions also aimed to explore how the intervention affected students' self-regulation and critical thinking. Moreover, the questions were designed to elicit information regarding effectiveness of the self-assessment proforma from the students' perspective. Students in these interviews were free to use either their mother tongue, Arabic, or English, and each interview was taped and lasted about 15 to 20 minutes.

4.6.3. Recruitment of the participants

4.6.3.1. Recruiting the participants for the first phase of the research (intervention and pre and post surveys)

As discussed in section 4.5, I contacted the dean and vice dean of the ELI at this Saudi HEI to attain permission to conduct the study. It should be noted that due to the Islamic and cultural beliefs of the country, the Saudi educational system is segregated by gender (Baki, 2004). Thus, as a female researcher, I randomly selected a section for this research that included 35 female students who were all aged 18 years old and over. All 35 students from the selected section agreed and appeared to be enthusiastic about taking part in the research. Ultimately, however, the participant number decreased to 27 students, as 4 students dropped the English course after the pre-test, and a further four students withdrew from the study during the intervention.

4.6.3.2. Recruiting the participants for the second phase of the study (assessment proformas, recording, and interviews)

At the end of the intervention and the post-test, the students' first and last recordings of their speaking activities and their self-assessment proformas were collected. However, although all students had already signed the consent form and agreed to provide their recording and proformas and agreed to take part in interviews, some students gave some verbal and nonverbal signs of discomfort about sharing their recordings or being recorded during the interview. Therefore, the participants of the first phase were further asked, via another consent form that included a separate point for the interview, if they consented to being approached for the study's semi-structured interview. Unfortunately, only 10 of the students who participated in the first phase agreed to this and signed the second consent form. The recordings and self-assessment proformas were collected from the students, and those 10 students were interviewed.

4.6.4. Validity and reliability

It is critical to generate accurate and trustworthy information and results in any kind of research (Merriam & Tisdell, 2015). However, what these ideas mean in practice is intensely contested. According to Jupp (2006), validity refers to the degree to which an indicator or variable accurately measures the theoretical idea which it is intended to measure, while reliability refers to the degree to which a measuring device consistently

produces accurate results. Within mixed methods research, in particular, it is vital that each method supports the other methods involved in the research, complementing other methods' weaknesses through its strengths (Onwuegbuzie & Johnson, 2006). Moreover, validity and reliability could mainly imply standard conditions of quality or quality criteria, which is also known variously as credibility, rigour, or trustworthiness (Morrow, 2005). However, to determine the quality of a research in any paradigm, the research should be assessed in terms of the paradigm's specific terminology (Healy & Perry, 2000). Due to the interpretive character of this study, conventional definitions of validity and reliability may be regarded as less relevant than they would be in quantitative research. Nonetheless, the results of any research must be acceptable, reasonable, and trustworthy (Lincoln & Guba, 1985). Thus, this research employs Lincoln and Guba's (1985) postpositivist method to evaluating the quality of interpretative research.

In their postpositivist method, Lincoln and Guba (1985) posit that the standard conditions of quality or the trustworthiness of a research study is critical when determining the study's value. According to this method, determining the trustworthiness of an interpretive research involves four evaluative criteria: credibility, transferability, dependability, and confirmability (Lincoln & Guba, 1985). These four criteria are parallel to the four criteria rooted in methodological positivism: internal validity, external validity and generalisability, reliability, and objectivity (Yanow & Schwartz-Shea, 2015) (see Table 4-3 for details). All four of these quality criteria were defined by Lincoln and Guba (1985). Credibility is interpreted as the degree of trust a reader can have in the truth of the study results; it determines if the study results reflect credible information derived from the participants' original data and offer an accurate interpretation of the participants' original perceptions (Lincoln & Guba, 1985). Transferability refers to the extent to which the findings of a study may be applied to different contexts or settings with other participants. Dependability refers to the stability of results over time, and confirmability refers to the extent to which the research study's conclusions can be verified by other researchers (Lincoln & Guba, 1985).

Table 4-3 Interpretive approaches to evaluative criteria (Yanow & Schwartz-Shea, 2015)

Criterion	Terms Used in Methodological Positivism	Lincoln and Guba (1985): Parallel terms
Truth value	Internal validity	Credibility
Applicability	External validity/ Generalisability	Transferability
Consistency	Reliability	Dependability
Neutrality	Objectivity	Confirmability

Moreover, Lincoln and Guba (1985) offered a set of techniques for conducting research that leads to a study being trustworthy: triangulation, prolonged engagement, dependability, and confirmability audits (for more details, see

Table 4-4 Lincoln and Guba's (1985) techniques to achieve *trustworthiness*). In this research, some of these strategies have been applied and are discussed in further detail in this section; these strategies help to establish validity and reliability, ensuring that the present study's results give an adequate measure of EFL learners' perceptions of the impact of self-assessment on their critical thinking, self-regulated learning, and performance.

Table 4-4 Lincoln and Guba's (1985) techniques to achieve trustworthiness

Criterion	Strategy	Definition
Credibility	Prolonged engagement	Spending an adequate amount of time in the field to learn or comprehend the culture, social environment, or phenomenon of interest.
	Persistent observation	Identifying and concentrating on the features and aspects of a situation that are most pertinent to the problem or issue at hand.
	Triangulation	Using various data sources in a research in order to gain more information.
	Peer debriefing	Exposing oneself to a disinterested peer in a way similar to an analytical session with the goal of examining elements of the inquiry that may otherwise stay merely implicit in the inquirer's thinking.
	Negative case analysis	Refining an analysis until it can account for or explain the majority of cases.
	Referential adequacy	Recognising a subset of data that should be preserved but not analysed.
	Member checks	Validating data, analytic categories, interpretations, and conclusions with members of the original data-gathering groups.
Transferability	Tick description	After thoroughly investigating a phenomena, assessing the degree to which the findings obtained are transferrable to different periods, places, circumstances, and individuals.
Dependability	Triangulation	Using various data sources in a research in order to gain more information.
	Dependability audits	Examining the investigation's method (how data were gathered, stored, and the data's accuracy)
Confirmability	Confirmability audits	Conducting an examination of the product to ensure that the findings, interpretations, and recommendations are backed up by data.
	Audits trail	Producing a detailed description of the processes involved in conducting research, from the inception of the research through the reporting of results.
	Triangulation	Using various data sources in a research in order to gain more information.

Moreover, in this section I explore validity and reliability threats that should be considered in the quantitative and qualitative components of mixed methods research. The mixed methods approach was selected for this study as a triangulation strategy. In other words, the mixed methods approach assists in bringing various views from various sources of data in order to improve methodological rigour; the use of several techniques allows researchers to more effectively present a credible narrative of reality (Calfee & Sperling, 2010). Using only one method for data collection could present a limited perspective of an issue; in contrast, using a combination of methods may provide for a valid and comprehensive understanding of the issue (Calfee & Sperling, 2010; Creswell & Creswell, 2017; Johnson & Christensen, 2004). Additionally, triangulation in its many forms has been deemed beneficial for increasing the dependability/reliability of any research (Lillis, 2006; Lukka, 1988). Triangulation in this research was undertaken through the employment of different tools for collecting the data, namely the pre- and post-tests, self-assessment proforma, audio recordings, and interviews. The pre- and post-tests, self-assessment proforma, and audio recordings supported the interviews and were all conducted prior to the interviews. Triangulating via these methods also supported the dependability of the findings by gaining saturation of the data in the study (Stavros & Westberg, 2009).

In the quantitative part of this research, the validity and reliability of the pre- and post-tests were confirmed since they were not constructed by the researcher, but were specifically designed to explore and measure SRL and critical thinking. The SRL-SRS has been proven to be a valid and reliable instrument for assessing SRL in the work of Toering et al. (2012) and Lucieer et al. (2016). Toering et al.'s (2012) work indicated that the SRL-SRS is a reliable instrument to measure SRL as a relatively stable attribute. The reliability coefficients, which vary from 0.00 to 1.00, assist in estimating the level of error related with the surveys. Following the basic rules given by the United States Department of Labor (1999) for evaluating a reliability coefficient, anything above 0.89 is deemed "exceptional", within 0.80–0.89 is considered "acceptable", within 0.70–0.79 is regarded "sufficient", and less than 0.70 is deemed "inadequate" and may have a narrow range of applicability (Saad et al., 1999). The reliability and internal consistency of the SRL-SRS questionnaire was checked and proven by Toering et al. (2012). The Cronbach's alpha for each component of the questionnaire was as follows: planning=0.81, self-monitoring=0.73, self-evaluation=0.82, reflection=0.78, effort=0.85, and self-efficacy=0.81. Toering et al.'s (2012) work also provided support for the validity of the instrument through confirmatory

factor analysis. The SRL-SRS has been used in multiple studies and contexts, having been applied to examine the impact of SRL strategies and academic achievements on recruitment of graduates in the KSA (Almutairi & Hasanat, 2018) and the development of SRL at medical school's pre-clinical stage (Lucieer et al., 2016). In the present research, the reliability and internal consistency of the data collected through the SRL-SRS questionnaire had to be checked using Cronbach's alpha coefficient. The Cronbach's alpha for each component of the questionnaire was as follows: planning=0.75, self-monitoring=0.74, self-evaluation=0.74, reflection=0.74, effort=0.84, and self-efficacy=0.76. Thus, the reliability of the SRL-SRS should be adequate.

Moreover, in order to increase the reliability of the questionnaire, back-translation and consultation were used to ensure good translation, as suggested by Filep (2009). Accordingly, the English questionnaire was first translated into Arabic then back-translated into English. Then, five bilingual language instructors were consulted and asked to check the clarity of the translated survey. No modifications were needed to the survey based on the instructors' feedback.

The validity and reliability of the Watson-Glaser II Critical Thinking Appraisal has also been proven in several studies. The test manual contains sufficient evidence of face, content, criteria, and construct validity (Watson & Glaser, 2010). The manual also reports that the Watson-Glaser II has been shown to be sufficiently reliable over time through test-retest reliability, and internal consistency has been proven through Cronbach's alpha coefficient (Watson & Glaser, 2010). Hassan and Madhum (2007) noted that the Watson-Glaser II is distinguished by its long history of use to measure critical thinking in different countries and diversified contexts and fields. Some examples of its use in the last 10 years include the comparison of critical thinking skills between special education school students and their graduates (Zascavage et al., 2007); the measurement of critical thinking of business administration students (Coleman et al., 2012); the evaluation of Middle Eastern students' critical thinking (Alshraideh, 2015) and the evaluation of US undergraduates (Obias, 2015); the evaluation of the relationship between academic achievement, self-regulation, and critical thinking in HE (Ghanizadeh, 2017); the measurement of Iranian students who took English language classes in terms of their critical thinking (Hashemi & Zabihi, 2012); the analysis of the relationship between critical thinking and personal reputation (Sandhu & Sharma, 2015); the measurement of the reading ability of students as measured by the TOEFL (Fahim et al., 2010); and the evaluation of the critical abilities of instructors of English as a foreign language in relation to their level of professional

achievement (Birjandi & Bagherkazemi, 2010). From the aforementioned, it can be ascertained that the Watson-Glaser II has been applied in a variety of settings around the globe with a wide variety of demographic groups; however, the majority of its applications have been with teens, undergraduates, and workers. This adds to the validity and reliability of the test.

As for speaking scores' reliability, providing clear and explicit instructions for scoring increases reliability as noted by Hughes (2003). Language tests or scores are also considered to be reliable when the tests or scores are dependable and consistent, as noted by Brown and Abeywickrama (2010). Accordingly, scoring was done using a detailed rubric provided by the institution, and I rated the audio recordings twice with a duration of a month between the two ratings and two sets of scores. Then, the reliability of the scores was determined through Cronbach's alpha coefficient, which revealed that the reliability was exceptional in the first and last recordings: Cronbach's alpha value was 0.96 for the first recording and 0.95 for the last recording, both well above 0.89 (see Table 4-5).

Table 4-5 1st Reliability statistics of speaking task

	Cronbach's Alpha	**No of Items**
1st audio recording	**.97**	2
Last audio recording	**.95**	2

As for speaking scores' validity, the two sets of scores should be compared to the learners' actual or true scores as given by the teacher according to Borman et al. (2001). Accordingly, the Cronbach's alpha coefficient was used to determine the rating accuracy and dependability between the three sets of scores. Table 4-6 reveals that the reliability is deemed exceptional between the three sets of scores, with Cronbach's alpha value at 0.95 and 0.93.

Table 4-6 2nd Reliability statistics of speaking task

	Cronbach's Alpha	**No of Items**
1st audio recording	**.95**	3
Last audio recording	**.94**	3

In the qualitative part of the research, the standard conditions of quality, which may refer to validity and reliability, are credibility, dependability, and transferability (Lincoln & Guba, 1985). Credibility is critical in both qualitative and quantitative research; in other words, regardless of method, it is critical to assess and demonstrate the study's credibility (Golafshani, 2003). According to Patton (2015), the credibility of quantitative research is reliant upon the instrument design; in qualitative research, in comparison, the researcher is the instrument. The researcher is, therefore, significant in achieving credibility in qualitative research. Similar to Patton (2015), Golafshani (2003) suggested that, in qualitative research, credibility is decided by the ability and effort of the researcher. This suggestion is applicable to the qualitative part of the present research, as I put significant effort into interacting with the participants throughout the data collection process in order to acquire a rich source of information and deeper analysis. In order to increase the reliability of the qualitative part of the research further, back-translation, consultation, and piloting were used to ensure good translation, as suggested by Filep (2009). I invited five bilingual language instructors to check the clarity of the self-assessment proforma and the interview questions in both versions, English and Arabic. No modifications were needed to the survey based on the instructors' feedback. Additionally, audio recording equipment was used during the interviews to provide an evidence base that could ensure that all details from the interviews with participants would be available for analysis (Gay et al., 2009). The findings' statistical generalisability may be regarded as limited due to the fact that the study was performed at a single institution; however, as previously discussed, this research was concerned with transferability that is concerned with investigating a single case and strove to comprehend and contextualise a unique situation (Gialdino & Vasilachis, 2009). According to Trochim (2005), the transferability of study results may be determined by comparisons between the location, individuals, and context; thus, given the centralised nature of the Saudi education system, one might argue that the results are also relevant beyond the study's particular setting to other HEIs in the KSA.

4.7. Conclusion

This chapter aimed to address the research philosophy and design of this research. It also aimed to discuss the sample and the tools used to serve the aim of this research, which was to explore the EFL learners' perceptions on the impact of self-assessment on their self-regulated learning, critical thinking, and achievement. The chapter first addressed the interpretive paradigm adopted in this research and the rationale for employing a mixed method research design. Mixed method research design was the most appropriate design to

obtain rich data in the context of the research, namely Saudi Arabia. The chapter then discussed the sample of the study and the instruments applied for data collection, which included pre- and post-questionnaires, audio recordings, and interviews. Additionally, the different means to assure the validity and the reliability of these instruments were discussed.

The following two chapters, Chapters 5 and 6, present the methods used to analyse the data collected in this study. The findings from the three research instruments in relation to students' perceptions and experience of the effect of self-assessment on EFL learners' self-regulated learning, critical thinking, and achievement are also discussed in Chapters 5 and 6.

Chapter 5 : Quantitative Data Analysis

5.1. Introduction

Chapter 4 outlined the research methodology for this study, including the use of the mixed-methods approach to explore and understand EFL learners' lived experiences and perspectives of the impact of self-assessment on their self-regulated learning, critical thinking, and achievement in English speaking classrooms in Saudi higher education institutions (HEIs) through pre- and post-surveys, audio-recording, self-assessment proformas, and semi-structured interviews. Both Chapters 5 and 6 outline the methods of analysis and present the findings of this research, and Chapter 7 then discusses the findings in relation to the literature. Chapter 5 focuses on the quantitative data collected in this study, and Chapter 6 then focuses on the qualitative data. The order of the presentation of the data is based in part on the chronological order in which data were collected. Moreover, I believe that it is necessary to move from the broader view to the individual. Therefore, after this chapter discusses first the broader view based on the quantitative data on students' perception prior to and after the intervention, Chapter 6 moves on to explore individual views through the qualitative data.

Chapter 5 addresses the methods used in the analysis of the quantitative data collected in this study, namely the pre- and post- surveys that addressed EFL students' perceptions regarding their self-regulated learning strategies, and measured the students' critical thinking prior to and after the intervention. The chapter also reports the findings and results of both the descriptive and inferential statistics. The discussion of the process of analysing the quantitative data and the presentation of the findings are integrated and presented under three sections (5.2, 5.3, and 5.4), each of which comprises several subsections.

5.2. EFL learners' perceptions of their self-regulated learning

As explained in Chapter 4, the Self-Regulation of Learning Self-Report Scale (SRL-SRS) was used to collect participants' perception of their self-regulatory skills before and after the intervention; the total number of survey respondents was n=27. Data management and analysis were performed using SPSS software (Version 26). The SRL-SRS employed the Likert scale, an ordinal scale of measurement used in ordering categories (Cohen et al., 2018).

The participants' responses for all six sections in the SRL-SRS were first analysed using descriptive statistical methods. The frequency and percentages for each statement were calculated based on the ratings provided by participants. Inferential statistics were also used to determine if there were statistically significant variations in learners' perceptions of their self-regulated learning prior to and after intervention. Non-parametric tests were used, as non-parametric approaches are generally optimal for analysing data on ordinal scales of measurements (Cohen et al., 2018; Pallant, 2010; Pallant, 2016). The use of parametric techniques procedures would have been unsuitable, as the data were not normally distributed; thus, the normality assumption for using parametric techniques with ordinal scale was not met. In comparison, non-parametric tests do not make assumptions about distribution. Therefore, non-parametric tests were used to analyse the survey. The Wilcoxon Signed Rank Test, which is the equivalent of Paired Sample T-test, was used in this research in order to determine whether there were statistically significant differences between two related samples, namely learners' perceptions in the pre- and post-surveys. However, statistical significance is proportional to the sample size and only intended to inform the researcher whether or not a certain outcome, such as a difference or a correlation, happened by chance (Cohen et al., 2018; Pallant, 2010; Pallant, 2016). Thus, it is argued that a measure of effect size, rather than statistical significance, may be more informative as statistical significance alone does not imply effect, and impact is what most researchers pursue (Cohen et al., 2018). Therefore, after determining effect size (r), it was necessary to use Equation 1 to measure the effect size or the strength of association and to determine effect size using Cohen's (1988) criteria of strength of association, namely .1 = small effect, .3= medium effect, .5= large effect (Pallant, 2010).

$$r = \frac{z}{\sqrt{N}}\ , \qquad \textit{Equation 1}$$

where N = total number of cases, and the z value is generated using Wilcoxon's Signed Rank Test.

The six sections of the SRL-SRS survey consisted of planning, self-monitoring, evaluation, reflection, effort, and self-efficacy. Each section evaluated a number of variables in order to target relevant learner perceptions. For example, the planning section consisted of nine statements incorporating variables that concerned learners' perceptions of their planning skills. Furthermore, because this research was interested in all aspects of learners' perceptions of their planning, it was necessary to not only analyse the responses to all these variables but also create a new variable that could represent planning, for example, through SPSS. This process, referred to as transforming variables, entails transforming data from a group of categorical variables into a single new continuous variable (UoS, 2014). In order

to transform these variables, SPSS collapsed each learners' responses for each statement in the section into one total score to calculate the average. For example, in the planning section, a learner who answered "4", or "almost always", in response to all nine statements on a four-point Likert scale of the planning section would score 36, as 4+4+4+4+4+4+4+4+4= 36, on the new scale for the new variable. That new scale ranges from 4 to 36, given that nine statements were in the section. The average of this score represents the score of the planning section for that learner.

Both descriptive and inferential statistics of the SRL-SRS prior to and after the intervention are discussed in the following sub-sections.

5.2.1. EFL learners' perceptions of their planning skill in the SRL-SRS survey

This section reports results from the first section of the SRL-SRS survey, which aimed to explore to what extent EFL learners believe they can perform a certain ability or task in relation to their planning skills and process. The learners marked their perception on a four-point Likert scale; the higher the number chosen, the more the participant believed that they could perform a certain skill or ability. This section contained nine statements. Table 5-1 presents the descriptive statistics from both pre- and post-surveys, including the frequency and percentages of the full sample (n=27) for each statement.

Table 5-1 Descriptive statistics for the full sample on perceptions regarding planning on the pre- and post-SRL-SRS survey for all statements

Statements	N		Pre-survey					Post-survey				
			Almost never	sometimes	often	Almost Always	Total	Almost never	sometimes	often	Almost Always	Total
1. I determine how to solve a problem before I begin.	27	F	0	9	12	6	27	0	1	8	18	27
		%	0%	33.3%	44.4%	22.2%	100%	0%	3.7%	29.6%	66.7%	100%
2. I think through in my mind the steps of a plan I have to follow.	27	F	2	7	12	6	27	0	1	8	18	27
		%	7.4%	25.9%	44.4%	22.2%	100%	0%	3.7%	29.6%	66.7%	100%
3. I try to understand the goal of a task before I attempt to answer.	27	F	0	7	10	10	27	0	1	9	17	27
		%	0%	25.9%	37.0%	37.0%	100%	0%	3.7%	33.3%	63.0%	100%
4. I ask myself questions about what a problem requires me to do to solve it, before I do it.	27	F	2	10	8	7	27	0	2	13	12	27
		%	7.4%	37.0%	29.6%	25.9%	100%	0%	7.4%	48.1%	44.4%	100%
5. I imagine the parts of a problem I still have to complete	27	F	3	11	8	5	27	0	3	12	12	27
		%	11.1%	40.7%	29.6%	18.5%	100%	0%	11.1%	44.4%	44.4%	100%
6. I figure out my goals and what I need to do to accomplish them.	27	F	0	6	10	11	27	0	2	9	16	27
		%	0%	22.2%	37.0%	40.7%	100%	0%	7.4%	33.3%	59.3%	100%
7. I carefully plan my course of action to solve a problem.	27	F	1	8	11	7	27	0	3	8	16	27
		%	3.7%	29.6%	40.7%	25.9%	100%	0%	11.1%	29.6%	59.3%	100%
8. I clearly plan my course of action to solve a problem.	27	F	1	8	11	7	27	1	3	8	15	27
		%	3.7%	29.6%	40.7%	25.9%	100%	3.7%	11.1%	29.6%	55.6%	100%
9. I develop a plan for the solution of a problem.	27	F	3	6	11	7	27	1	3	7	16	27
		%	11.1%	22.2%	40.7%	25.9%	100%	3.7%	11.1%	25.9%	59.3%	100%

(F: Frequency)

Overall, the findings of the data presented in Table 5-1 reveal significant changes in the learners' perceptions of different planning processes after the intervention. The first four statements in the survey aimed to explore learners' perceptions of their planning strategies before starting a task. Regarding the first two statements in the pre-survey, two thirds of the learners, or 44.4% (12 participants) often and 22.2% (six participants) almost always, felt that they tried to determine how to solve a problem before beginning and thought through the steps that they would have to follow. In the post-survey, the number of participants who felt this way increased: the majority of the participants (26 participants), or 29.6% (eight participants) often and 66.7% (18 participants) almost always, planned in this way. One possible reason for this increase may be that, through the intervention, learners were given a way and an opportunity to identify the area on which they should focus while doing a task; therefore, the learners were prompted to think about and determine the steps needed for them to complete a task and succeed. Students' responses to the third statement suggest that more than two thirds of the participants, with 37.0% (10 participants) often and 37.0% (10 participants) almost always believing that, when performing tasks, they tried to understand the goal before attempting to answer. This percentage increase in the post-survey, with most participants feeling that they often, with 33.30% (nine participants), or almost always, with 63.0% (17 participants), made an effort to comprehend the goal of a task before trying to accomplish the task. In response to the fourth statement, more than half the participants, a total of 55.5%, felt that they often, with 29.6% (eight participants), or almost always, with 25.9% (seven participants) asked themselves what a problem required them to do to solve it before completing the problem. One possible reason for the positive increase could be that, through the intervention learners, were given a means and an opportunity to identify the area in which they should focus on while completing a task, meaning that they were able to think and determine the necessary steps for them to set a goal to complete a task and succeed.

Moreover, statements from four to nine aimed to explore learners' perceptions of their planning strategies while accomplishing a task. Students' responses to the fifth statement in the pre-survey reveal that almost half the participants felt that they often, with 29.6% (eight participants), or almost always, with 18.5% (five participants), thought of the components of an issue that they had yet to solve. However, the number of participants who felt this way increased after the intervention, and responses to the same statement in the post-survey reveal that more than 85% of the participants (24 participants) felt that they often, with 44.4% (12 participants), or almost always, with 44.4% (12 participants) thought

of the components of an issue that they have yet to solve. Table 5-1 revealed EFL learners' responses to statements number seven, eight, and nine in the pre survey indicated that two thirds of the participants (66.6%) felt that they often, with 40.7% (11 participants), or almost always, with 25.9% (seven participants) developed and planned their course of action to solve a problem carefully and clearly. However, the number of participants who felt this way increased after the intervention, and their responses to the same statements in the post-survey reveal that over 85% of the participants felt that their line of action to resolve an issue was carefully and clearly planned. Learners' responses to these statements may be attributed to the fact that, by raising learners' awareness of the criteria through self-assessment proformas, they were able to plan the means of completing the task carefully and clearly.

Furthermore, students' responses to all nine statements that aimed to explore their perceptions of their planning process were merged, as explained in Section 5.2, and presented as one single indicator to represent learners' perceptions of their planning processes prior to and after the intervention. Table 5-2 presents the descriptive statistic of the planning section as a whole from both pre- and post-surveys, including the mean of the full sample (n=27) for each statement. It is apparent from Table 5-2 that the mean of participants' responses in the planning section in the post-SRL-SRS survey is higher than those responses from the pre-survey, and the mean difference is 0.63.

However, in order to confirm whether the difference in means between the pre- and post-survey was statistically significant, Wilcoxon' Signed Rank Test was used, as discussed in Section 5.2.

Table 5-2 Descriptive statistics of the full sample scores of the planning section in the pre- and post-SRL-SRS survey

	N	Pre-survey					Post-survey				
		Mean	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
Planning	27	2.85	2.88	.49	1.78	3.67	3.48	3.55	.43	2.67	4.00

Wilcoxon's Signed Rank Test was chosen to establish a comparison between the two related sets of data as a mean to determine any statistically significant difference between the two. These results are presented in Table 5-3.

Table 5-3 Statistical difference of the panning section in the pre and post-SRL-SRS survey

Test Statistics	
	Post-planning – Pre-Planning
Z	-4.44[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

The findings from the Wilcoxon Signed Rank presented in Table 5-3 reveal a statistically significant difference between learners' perceptions of their planning processes prior to and after intervention. The test reveals a statistically significantly positive change in learners' views, Z= -4.44, $p < .001$, with a large effect size (r = .85). The median score of the perceptions of planning increased from pre-survey (*Md* = 2.88) to post-survey (*Md*= 3.55).

Learners' perceptions of planning before and after the intervention indicate highly positive attitudes towards the intervention. The participants believed that they were more engaged while completing the task and using the different strategies while planning for the task completion more frequently. This result may indicate a positive impact of self-assessment on learners' perceptions and use of their planning skills. However, it is not possible to be certain of the reasons for this positive perception before examining the qualitative data.

The findings and descriptive statistics from the second section in the survey, namely self-monitoring, are presented and discussed in the following sub-section.

5.2.2. EFL learners' perceptions of their self-monitoring skills in the SRL-SRS survey

This section presents the results of the second section in the SRL-SRS survey, which contained eight statements. The second section in the survey aimed to explore EFL learners' perceptions of their self-monitoring strategies. Learners indicated their perceptions on a four-point scale; the higher the number, the more confident that the learners were in their ability to exercise a certain skill. Table 5-4 presents the descriptive statistics from both the pre- and post-survey, including the frequency and percentages of the full sample (n=27) for each statement.

Table 5-4 Descriptive statistics of the full sample's perceptions of self-monitoring skills in the pre- and post-SRL-SRS survey

Statements	N		Pre-survey					Post-survey				
			Almost never	sometimes	often	Almost Always	Total	Almost never	sometimes	often	Almost Always	Total
1. While doing a task, I ask myself questions to stay on track	27	F	3	6	10	8	27	1	1	7	18	27
		%	11.1%	22.2%	37.0%	29.6%	100%	3.7%	3.7%	25.9%	66.7%	100%
2. While doing a task, I ask myself, how well I am doing	27	F	3	6	11	7	27	0	1	13	13	27
		%	11.1%	22.2%	40.7%	25.9%	100%	0%	3.7%	48.1%	48.1%	100%
3. I check my work while doing it.	27	F	0	6	10	11	27	0	1	5	21	27
		%	0%	22.2%	37.0%	40.7%	100%	0%	3.7%	18.5%	77.8%	100%
4. I check how well I am doing when I solve a task.	27	F	0	6	10	11	27	0	1	5	21	27
		%	0%	22.2%	37.0%	40.7%	100%	0%	3.7%	18.5%	77.8%	100%
5. I check my accuracy as I progress through a task.	27	F	1	5	9	12	27	0	1	6	20	27
		%	3.7%	18.5%	33.3%	44.4%	100%	0%	3.7%	22.2%	74.1%	100%
6. I know how much of a task I have to complete.	27	F	2	4	6	15	27	0	0	7	20	27
		%	7.4%	14.8%	22.2%	55.6%	100%	0%	0%	25.9%	74.1%	100%
7. I judge the correctness of my work.	27	F	2	4	11	10	27	0	0	3	24	27
		%	7.4%	14.8%	40.7%	37.0%	100%	0%	0%	11.1%	88.9%	100%
8. I correct my errors	27	F	1	3	5	18	27	0	0	4	23	27
		%	3.7%	11.1%	18.5%	66.7%	100%	0%	0%	14.8%	85.2%	100%

(F: Frequency)

The results in Table 5-4 reveal substantial changes in learners' views of self-monitoring processes prior to and after intervention. The first six statements in this section intended to gather information on learners' views of their self-monitoring tactics while completing a task. Learners' responses to the first statement in the pre-survey indicate that two thirds of the participants believed that they asked themselves questions while completing a task as a means to stay on track. 37.0% of the participants (10 participants) believed that they often asked themselves, and 29.6% (eight participants) believed that they almost always used this technique to stay on track. Interestingly, students' responses in the post-survey show that almost all the participants used this technique to stay on track, with 66.7% (18 participants) indicating that they almost always used this technique when completing tasks, and 18.5% believing that they often do so. However, it should be noted that the second statement in the survey aimed to gather information on learners' views of whether they asked themselves to track how well they were performing the task, and the data reveal that, in the pre-survey, the participants gave almost identical responses to the first statement. 40.7% of the participants believed that they often asked themselves how well they were performing a task and 25.9% believed that they almost always used this technique to stay on track. However, although learners' responses to the post-survey show that, similar to the first statement, almost all participants believed that they asked themselves how well they were performing doing a task, only 48.1% of the participants (13 participants) believed that they almost always used this technique, and the other 48.1% of the participants (13 participants) believed that they often asked questions to assess their performance in a task. The change in attitudes toward these statements may be because the awareness of assessment criteria and requirement enable the students to set objectives and goals that they try to reach.

Moreover, the data presented in Table 5-4 for participants' responses for the third, fourth, and fifth statements showed that, prior to the intervention, over 75% of the participants felt that they checked their work and the accuracy of this work. Almost 37.0% of the learners believed that they often checked their work, and almost 40.7% felt that they almost always checked their work. Remarkably, the number of participants who felt this way increased after the intervention; their responses to the same statements in the post-survey suggested that almost all the participants felt that they ensured the quality of their work and its correctness, with almost 77.8% feeling that they almost always and almost 18.5% believed that they do so. Table 5-4 also shows learners' responses for the sixth statement in the survey, revealing that over 75% of the EFL learners believed that they were aware of how much of a task remained to be completed, with 22.2% (six participants) believing that they

were often aware and 55.5% (15 participants) believing that they were almost always aware of the amount of task in hand. Remarkably, learners' responses after the intervention suggest that almost all the participants felt that they were aware of the amount of task that remained to be complete. 25.9% (seven participants) of the learners believed that they were aware of what was needed to be completed, and 74.1% (20 participants) felt that they were almost always ware. The possible reason for this finding is that using the audio recordings enabled the learners to monitor their speaking performance as they had evidence of their performance, which would allow them to check and identify their mistakes and weaknesses.

Furthermore, students' responses to the seventh statement in the pre-survey reveal that over 75% of the participants felt that they often, with 40.7% (11 participants), or almost always, with 37.0% (10 participants), judged the correctness of their work after completing it. Surprisingly, the participants' responses to this question after the intervention suggest that all the participants felt that they were able to judge the correctness of their work. Only 11% (three participants) of the learners' stated that they often judged their work, while 88.9% (24 participants) of the learners stated that they almost always did so. The data in Table 5-4 presents learners' responses to the eighth statement, which aimed to explore to what extent learners corrected their errors when completing tasks. The data reveal that, in the pre-survey, approximately 85% of the participants stated that they either often or almost always corrected their errors. 18.5% (five participants) of the learners believed that they often corrected their work, and 66.7% (18 participants) of the learners' stated that they almost always did so. Remarkably, the number of participants who felt this way increased after the intervention and their responses to the same statement in the post-survey reveal that all the participants stated that they corrected their errors with 85.2% (23 participants) feeling that they almost always corrected their mistakes, and 14.8% (four participants) stating that they often did so.

Furthermore, learners' responses to all eight statements, which aimed to explore their perceptions of their self-monitoring processes, were merged and presented as one single indicator to represent learners' perceptions of their self-monitoring processes prior to and after the intervention. Table 5-5 presents the descriptive statistic of the self-monitoring section as a whole from both the pre- and post-survey, including the mean of the full sample (n=27) for each statement. It is evident from Table 5-5 that the mean of participants' responses in the self-monitoring section in the post survey is higher than their responses in the pre-survey, with a mean difference of 0.58.

Table 5-5 Descriptive statistics of the full sample scores of the self-monitoring section in the pre- and post-SRL-SRS survey

	N	Pre-survey					Post-survey				
		Mean	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
Self-Monitoring	27	3.11	3.25	.53	1.75	4.00	3.69	3.75	.31	2.75	4.00

In order to confirm whether the difference in means between learners' perceptions of their self-monitoring processes in the pre- and post-survey was statistically significant, Wilcoxon's Signed Rank Test was used, the findings of which are presented in Table 5-6.

Table 5-6 Statistical difference of the self-monitoring section in the pre and post-SRL-SRS survey

Test Statistics	
	Post-self-monitoring – Pre-self-monitoring
Z	-4.37[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-6 presents the results of Wilcoxon's Signed Rank test and indicates a statistically significant difference between learners' perceptions of their self-monitoring processes prior to and after the intervention. The test reveals a statistically significant positive change in learners' perceptions, Z= -4.37, $p < .001$, with a large effect size (r = .0.84). The median score on perceptions of self-monitoring increased from the pre-survey (*Md* = 3.25) to the post-survey (*Md*= 3.75).

In summary, learners' perceptions of the planning section before and after the intervention show high positive attitudes and confidence in their ability to monitor, check, and correct their performance at different stages, which may indicate the positive impact of self-assessment in this positive change of their perceptions of their self-monitoring skill and processes. The reasons for these changes in perceptions may be clearer when discussing these results side-by-side with the results of qualitative data.

The findings and descriptive statistics of the second section in the survey self-monitoring are presented and discussed in the following sub-section.

5.2.3. EFL learners' perceptions of their evaluation strategy in the SRL-SRS survey

This section reports the findings of the third section of the SRL-SRS survey. The third section in the survey included eight statements and aimed to explore EFL learners' perceptions of their self-monitoring strategies. The learners indicated their perceptions on five scales: the higher the scale, the more confident that they were in their ability to perform a certain skill. Table 5-7 presents the descriptive statistics from both pre- and post-survey, including the frequency and percentages of the full sample (n=27) for each statement.

Table 5-7 Descriptive statistics of the full sample's perceptions of their evaluation skills in the pre- and post-SRL-SRS survey

Statements	N		Pre-survey						Post-survey					
			Never	Rarely	sometimes	often	Always	Total	Never	Rarely	sometimes	often	Always	Total
1. I look back and check if what I did was right.	27	F	0	2	11	10	4	27	0	0	0	4	23	27
		%	0%	7.4%	40.7%	37.0%	14.8%	100%	0%	0%	0%	14.8%	85.2%	100%
2. I double-check to make sure I did it right.	27	F	1	3	12	9	2	27	0	0	1	12	14	27
		%	3.7%	11.1%	44.4%	33.3%	7.4%	100%	0%	0%	3.7%	44.4%	51.9%	100%
3. I check to see if my calculations are correct.	27	F	1	3	4	12	7	27	0	0	2	11	14	27
		%	3.7%	11.1%	14.8%	44.4%	25.9%	100%	0%	0%	7.4%	40.7%	51.9%	100%
4. I look back to see if I did the correct procedures.	27	F	0	1	6	10	10	27	0	0	0	6	21	27
		%	0%	3.7%	22.2%	37.0%	37.0%	100%	0%	0%	0%	22.2%	77.8%	100%
5. I check my work all the way through the problem	27	F	0	1	8	9	9	27	0	0	1	7	19	27
		%	0%	3.7%	29.6%	33.3%	33.3%	100%	0%	0%	3.7%	25.9%	70.4%	100%
6. I look back at the problem to see if my answer makes sense.	27	F	0	2	7	11	7	27	0	0	1	5	21	27
		%	0%	7.4%	25.9%	40.7%	25.9%	100%	0%	0%	3.7%	18.5%	77.8%	100%
7. I stop and rethink a step I have already done.	27	F	0	1	1	10	15	27	0	0	3	5	19	27
		%	0%	3.7%	3.7%	37.0%	55.6%	100%	0%	0%	11.1%	18.5%	70.4%	100%
8. I make sure I complete each step	27	F	0	1	3	5	18	27	0	0	1	3	23	27
		%	0%	3.7%	11.1%	18.5%	66.7%	100%	0%	0%	3.7%	11.1%	85.2%	100%

Generally, the data in Table 5-7 suggest remarkable changes in learners' perceptions of their evaluation processes pre- and post-intervention. Students' responses to the first statement in the pre-survey reveal that approximately half of the participants stated that they checked whether or not what they did was correct, with 37.0% (10 participants) stating that they often checked what they had done after they finished. 14.8% (four participants) of the learners stated that they always did so. Surprisingly, the participants' responses to the same statement after the intervention reveal that all the participants felt that they evaluated and checked their work after finishing. 85.2% (23 participants) of the participants stated that they always did so, and only 14.8% (four participants) of the learners state that they often did so. Similarly, the data on the second statement in the survey revealed that 40.7% of the participants stated that they double-checked if what they did was right after finishing. 33.3% (nine participants) of the learners stated that they often re-evaluated their work, while only 7.4% (two participants) of learners stated that they always did so. Remarkably, the number of participants who stated that they double-checked their work increased after the intervention; their responses in the post-survey suggest that almost all of the participants stated that they double-checked and re-evaluate their work to determine if what they did was correct, with 85.2% (23 participants) feeling they always corrected their mistakes and 14.8% (four participants) stating that they often did so.

Moreover, Table 5-7 shows learners' responses to the third and fourth statements, which aimed to gather information on learners' evaluation processes in the details of the work. The data reveal that approximately 70% of EFL learners stated that they checked if the details of their work were correct and if they followed the correct procedures or format. The details of the responses on the third statement reveal that 44.4% (12 participants) of learners stated that they often checked if the details of their work were correct, and 25.9% (seven participants) of the learners stated that they always checked their work after finishing to check the details. Additionally, the findings of the fourth statement reveal that 37.0% (10 participants) of the learners stated that they often checked if they followed the correct procedures or format, and 37.0% (10 participants) stated that they always did so. Remarkably, the number of participants who stated that they often or always checked if the details of their work were correct, and if they followed the correct procedures or format, increased after the intervention and the responses in the post survey suggest that almost all the participants stated that they did so. The fifth statement in the questionnaire aimed to explore to what extent students checked if their answers made sense in relation to the

speaking tasks. Table 5-7 shows that, in the pre-survey, approximately two thirds of the participants stated that they checked their responses. In the details, 40.7% (11 participants) of learners stated that they often checked their answers, and 25.9% (seven participants) stated that they always did so. Surprisingly, learners' responses in the post-survey reveal that students tended to focus more on evaluating the meaning of their answers in relation to the speaking topics; 96.3% of the participants stated that they either often or always checked if their answers made sense. Table 5-7 shows that, in the post-survey, 77.8% (21 participants) of the participants, in comparison to 25.9% (seven participants) in the pre-survey, stated that they always checked if their answers make sense. Additionally, 18% (five participants) of the learners in the post-survey stated that they often checked their responses.

Furthermore, students' responses to the seventh statement in the pre- and post-survey reveal that, after the intervention, more learners tended to stop and reconsider previous steps when working on a task. In details, the data presented in Table 5-7 showed that, in the post-survey, 70.4% (19 participants) of the participants, in comparison to 55.6% (15 participants) in the pre-survey stated that they always stopped and rethought a step that they had already completed. Additionally, 18% (five participants) of learners in the post-survey stated that they stopped and rethought a step that they had already done. Table 5-7 presented learners' responses to the eighth statement, the data of which reveal that, in the pre-survey, 85% of the participants stated that they made sure that they completed each step. 18.5% (five participants) made sure that they completed each step, and 66.7% (18 participants) of the learners stated that they almost always did so. Remarkably, the number of participants who felt this way increased after the intervention, and their responses to the same statement in the post survey reveal that almost all the participants stated that they made sure that they completed each step, with 85.2% (23 participants) feeling they always corrected their mistakes and 14.8% (four participants) stating that they often did so.

Moreover, descriptive statistics of learners' responses to all eight statements, presented as one single indicator to represent learners' perceptions of their self-monitoring processes prior to and after the intervention, are presented in Table 5-8. Similar to the first two sections of the survey, it is evident from Table 5-8 that the mean of participants' responses in the evaluation section in the post survey is higher than responses in the pre-survey, and the mean difference is 0.74.

Table 5-8 Descriptive statistics of the full sample scores of the evaluation section in the pre- and post SRL-SRS survey

	N	Pre-survey					Post-survey				
		Mean	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
Evaluation	27	3.93	4.00	.53	2.50	4.75	4.67	4.75	.25	4.00	5.00

Moreover, Wilcoxon's Signed Rank Test was then used to investigate whether the difference in means between the pre- and post-survey was statistically significant or not. These findings are presented in Table 5-9.

Table 5-9 Statistical difference of the evaluation section in the pre and post-SRL-SRS survey

Test Statistics	
	Post-Evaluation – Pre-Evaluation
Z	-4.55[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

The Wilcoxon's Signed Rank test revealed a statistically significant difference between learners' perceptions of their evaluation processes prior to and after intervention. The test revealed a statistically significantly positive change in learners' perceptions, $Z = -4.55$, $p < .001$, with a large effect size ($r = .87$). The median score on the perceptions of their evaluation processes increased from pre- ($Md = 4.00$) to post-survey ($Md = 4.75$).

In summary, the participants' perceptions reveal higher willingness and engagement in evaluating their work at each stage while completing their task, until evaluating their performance in the audio recording. One possible explanation for these results is that learners were more aware of the requirement of the tasks and the assessment criteria, which could motivate them to work hard and evaluate this work to succeed.

The findings and descriptive statistics of the second section in the survey self-monitoring are presented and discussed in the following sub-section.

5.2.4. EFL learners' perceptions of their reflection strategy in the SRL-SRS survey

This section reports findings from the fourth section of the SRL-SRS survey. This section included five statements that sought to elicit information on learners' views of their reflection processes. The learners indicated their views and level of agreement with a five-point Likert scale ranging from "strongly disagree" to "strongly agree". Table 5-10 presents the descriptive statistics from both pre- and post-survey, including the frequency of agreement and percentages of the full sample (n=27) for each statement.

Table 5-10 Descriptive statistics of the full sample's perceptions of their reflection skills in the pre- and post SRL-SRS survey

Statements	N		Pre-survey						Post-survey					
			SD	D	N	A	SA	Total	SD	D	N	A	SA	Total
1. I reappraise my experiences so I can learn from them	27	F	1	1	13	6	6	27	0	0	1	8	18	27
		%	3.7%	3.7%	48.1%	22.2%	22.2%	100%	0%	0%	3.7%	29.6%	66.7%	100%
2. I try to think about my strengths and weaknesses.	27	F	0	2	8	9	8	27	0	1	0	7	19	27
		%	0%	7.4%	29.6%	33.3%	29.6%	100%	0%	3.7%	0%	25.9%	70.4%	100%
3. I think about my actions to see whether I can improve them	27	F	0	0	6	14	7	27	0	0	2	3	22	27
		%	0%	0%	22.2%	51.9%	25.9%	100%	0%	0%	7.4%	11.1%	81.5%	100%
4. I think about my past experiences to understand new ideas	27	F	0	3	6	12	6	27	0	0	0	6	21	27
		%	0%	11.1%	22.2%	44.4%	22.2%	100%	0%	0%	0%	22.2%	77.8%	100%
5. I try to think about how I can do things better next time	27	F	1	0	4	6	16	27	0	0	0	4	23	27
		%	3.7%	0%	14.8%	22.2%	59.3%	100%	0%	0%	0%	14.8%	85.2%	100%

(**SD**: Strongly Disagree, **D**: Disagree, **N**: Neither, **A**: Agree, **SA**: Strongly Agree, **F**: Frequency)

The data in Table 5-10 shows significant change in learners' perceptions of their reflection processes prior to and after the intervention. Remarkably, learners' responses to the first statement in this section in the pre- and post-survey reveal that, after the intervention, more learners agreed that they carefully considered their experiences in order to gain knowledge from them. Table 5-10 reveals that, in the post-survey, a total of 96.3% (26 participants) of the participants agreed that they reappraised their experiences to learn from them, with 66.7% (18 participants) strongly agreeing and 29.6% (eight participants) agreeing, in comparison to 44.4% (12 participants) in the pre-survey with 22.2% (six participants) strongly agreeing and 22.2% (six participants) agreeing. Table 5-10 also presents learners' responses to the second statements which reveal that the number of participants who agreed that they tried to weigh their strengths and weaknesses increased after the intervention. In details, the data reveals that, in the post-survey, a total of 96.3% (26 participants) of the participants agreed that they tried to weigh their strengths and weaknesses with 70.4% (19 participants) strongly agreeing and 25.9% (seven participants) agreeing in comparison to 62.9% (17 participants) in the pre-survey with 29.6% (eight participants) strongly agreeing and 33.3% (nine participants) agreeing.

Moreover, Table 5-10 presents learners' responses to the third statement. The data reveal that, in the pre-survey, 77.8% of the participants agreed that they re-thought their actions to see whether they could improve, with 25.9% (seven participants) strongly agreeing and 51.9% (14 participants) agreeing. Interestingly, the number of participants who agreed with this statement increased after the intervention in the post-survey, with a total agreement of 92.6% of the participants, with 81.5% (22 participants) strongly agreeing and 11.1% (three participants) agreeing. Surprisingly, learners' responses to the fourth statement in the pre- and post-surveys reveal that, after the intervention, more learners agreed that they thought about their past experiences to understand new ideas. In details, Table 5-10 reveals that, in the post-survey, all the participants (27 participants) agreed that they thought about their past experience to understand new ideas with 77.8% (21 participants) strongly agreeing and 22.2% (six participants) agreeing in comparison to 66.6% (18 participants) in the pre-survey with 22.2% (six participants) strongly agreeing and 44.4% (12 participants) agreeing. Finally, the responses to the fifth statement reveal that, in the pre-survey 81.5% of the participants (22 participants) agreed that they tried to think about how they could do things better next time, with 59.3% (16 participants) strongly agreeing and 22.2% (six participants) agreeing. Remarkably, the number of participants who agreed increased after the intervention and the responses in the post-survey reveal that all the participants tried to

think about how they could do things better next time, with 85.2% (23 participants) strongly agreeing and 14.8% (four participants) agreeing.

Furthermore, the descriptive statistic of learners' responses to all five statements, which are presented as a single indicator to represent learners' perceptions of their reflection processes prior to and after the intervention are presented in Table 5-11.

Table 5-11 Descriptive statistics of the full sample scores of the reflection section in the pre- and post SRL-SRS survey

	N	Pre-survey					Post-survey				
		Mean	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
Reflection	27	3.91	4.00	.65	2.00	5.00	4.72	4.80	.29	4.00	5.00

It is apparent from Table 5-12 that the mean of participants' responses in the evaluation section in the post-survey is higher than those in the pre-survey, with a mean difference of 0.81. Nonetheless, it was necessary to use Wilcoxon's Signed Rank Test to investigate whether the difference in means between the pre- and post-survey was statistically significant, the findings of which are presented in Table 5-12.

Table 5-12 Statistical difference of the reflection section in the pre and post-SRL-SRS survey

Test Statistics	
	Post-reflection – Pre- reflection
Z	-4.32[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test c. Based on negative ranks.	

Table 5-12 presents the results of Wilcoxon's Signed Rank test, revealing a statistically significant difference between learners' perceptions of their reflection processes prior to and after the intervention. The test indicates a statistically significantly positive change in learners' perceptions, $Z = -4.32$, $p < .001$, with a large effect size ($r = .83$). The median score on the perceptions of reflection increased from pre-survey ($Md = 4.00$) to post-survey ($Md = 4.80$).

Although the reasons for the high agreement by the learners after the intervention cannot be certain before comparing these findings to those from the qualitative data, it may be possible to suggest that one of the reasons is that the self-assessment engaged students in their learning and provided opportunity for them to reflect more, identify their strength and weaknesses, and think of means of improving and overcoming weaknesses and issues. The findings and descriptive statistics in the fifth section in the survey on effort are presented and discussed in the following sub-section.

5.2.5. EFL learners' perceptions of their effort in the SRL-SRS survey

This section reports the findings from the fifth section of the SRL-SRS survey. The fifth section of the survey included 10 statements and aimed to explore EFL learners' perceptions of the effort that they put into their work. The learners indicated their views on a five-point scale; the higher the scale, the more confident they are in the level of effort that they put into their work. Table 5-13 presents the descriptive statistics from the pre- and post-survey, including the frequency and percentages of the full sample (n=27) for each statement.

Table 5-13 Descriptive statistics of the full sample's perceptions of their effort in the pre- and post-SRL-SRS survey

Statements	N		Pre-survey					Post-survey				
			Almost never	sometimes	often	Almost Always	Total	Almost never	sometimes	often	Almost Always	Total
1. I keep working even on difficult tasks	27	F	1	4	11	11	27	0	0	9	18	27
		%	3.7%	14.8%	40.7%	40.7%	100%	0%	0%	33.3%	66.7%	100%
2. I put forth my best effort when performing tasks	27	F	1	3	4	19	27	0	1	5	21	27
		%	3.7%	11.1%	14.8%	70.4%	100%	0%	3.7%	18.5%	77.8%	100%
3. I concentrate fully when I do a task.	27	F	2	5	6	14	27	0	0	11	16	27
		%	7.4%	18.5%	22.2%	51.9%	100%	0%	0%	40.7%	59.3%	100%
4. I don't give up even if the task is hard.	27	F	1	6	11	9	27	0	3	9	15	27
		%	3.7%	22.2%	40.7%	33.3%	100%	0%	11.1%	33.3%	55.6%	100%
5. I work hard on a task even if it is not important	27	F	3	9	8	7	27	1	2	15	9	27
		%	11.1%	33.3%	29.6%	25.9%	100%	3.7%	7.4%	55.6%	33.3%	100%
6. I work as hard as possible on all tasks	27	F	0	6	8	13	27	0	1	12	14	27
		%	0%	22.2%	29.6%	48.1%	100%	0%	3.7%	44.4%	51.9%	100%
7. I work hard to do well even if I don't like a task.	27	F	3	10	8	6	27	1	1	17	8	27
		%	11.1%	37.0%	29.6%	22.2%	100%	3.7%	3.7%	63.0%	29.6%	100%
8. If I'm not really good at a task I can compensate for this by working hard.	27	F	0	10	11	6	27	0	3	11	13	27
		%	0%	37.0%	40.7%	22.2%	100%	0%	11.1%	40.7%	48.1%	100%
9. I am willing to do extra work on tasks in order to learn more.	27	F	2	1	9	15	27	0	1	4	22	27
		%	7.4%	3.7%	33.3%	55.6%	100%	0%	3.7%	14.8%	81.5%	100%
10. If I persist on a task, I'll eventually succeed.	27	F	0	2	7	18	27	0	0	5	22	27
		%	0%	7.4%	25.9%	66.7%	100%	0%	0%	18.5%	81.5%	100%

Generally, the data in Table 5-13 reveal changes in learners' perceptions of the level of effort that they put into their work prior to and after the intervention. Students' responses to the first statement in the pre-survey reveal that more than 80% of the participants believed that they were persistent in their efforts even when confronted with challenging tasks, with 40.7% (11 participants) stating that they often kept working even on difficult tasks, and 40.7% (11 participants) of the learners stated that they almost always did so. The participants' responses to the same statement after the intervention revealed that all the participants felt that they persisted in their efforts even when confronted with challenging tasks. 66.7% of the participants (18 participants) stated that they almost always did so, and 33.3% of the participants (nine participants) stated that they often did so. Table 5-13 also shows learners' responses to the second statements, and the data revealed that the number of participants who felt that, while completing assignments, they made their best effort increased after the intervention. In details, the data in Table 5-13 suggest that, in the post-survey, a total of 96.3% (26 participants) of the participants felt that they made a concerted effort while completing tasks, with 77.8% of the participants (21 participants) stating almost always and 18.5% (five participants) stating that they often felt this way in comparison to 85.2% (23 participants) in the pre-survey with 70.4% (19 participants) saying almost always and 14.8% (four participants) often felt that they put in their best effort when performing a task. Moreover, Table 5-13 presents learners' responses to the third statement, the data of which reveal that, in the pre-survey, 74.1% of the participants believed that, when working on a task, they gave it their all, with 51.9% (14 participants) saying almost always and 22.2% (six participants) often feeling that they concentrate fully when doing a task. Notably, in the post-survey all participants felt that they completely focused when working on a task, increasing to 59.3% (16 participants) of the participants almost always and 40.7% (11 participants) often feeling this way.

Moreover, statements from the fifth to eighth in this section aimed to gather information on how hard learners are working on tasks. Learners' responses to these statements in the pre- and post-survey reveal that, after the intervention, more learners felt that they worked hard when completing a task. In details, learners' responses to the fifth statement presented in Table 5-13 reveal that, in the post-survey, 88.9% of the participants (24 participants) felt that they devoted considerable effort to a task, regardless of its significance, with 33.3% (nine participants) stating almost always and 55.6% (15 participants) stating that they often felt this way in comparison to 55.5% (15 participants) in the pre-survey, with 25.9% (seven participants) almost always and 29.6% (eight participants) often feeling that they worked

hard on a task regardless of its importance. Furthermore, responses to the sixth statement reveal that, in the post-survey, 96.3% of the participants (26 participants) felt that they strived to give their best to all tasks, with 51.9%% of the participants (14 participants) stating almost always and 44.4% (12 participants) stating that they often felt this way in comparison to 78.3% (21 participants) in the pre-survey, with 48.1% (13 participants) stating almost always and 29.6% (eight participants) often feeling that they worked as hard as possible on all tasks. Similarly, responses to the seventh statement reveal that, in the post-survey, 92.6% of the participants (25 participants) felt that they strived to complete tasks successfully, even the task were unpleasant, with 29.6%% of the participants (eight participants) stating almost always and 63.0% (17 participants) stating that they often felt this way in comparison to 51.8% (14 participants) in the pre-survey, with 22.2% (six participants) stating almost always and 29.6% (eight participants) stating that they often felt that they worked hard to complete tasks effectively, even when they disliked a task. Furthermore, responses to the eighth statement revealed that in the post survey a total of 88.8% of the participants (24 participants) felt that, even if they felt that they were good at a task, they could compensate by working hard, with 48.1% of the participants (13 participants) stating almost always and 40.7% (11 participants) stating that they often felt this way in comparison to 62.9% (17 participants) in the pre-survey, with 22.2% (six participants) stating almost always and 40.7% (11 participants) stating that they often felt that, if they were not good at a task, they could compensate by working hard.

Furthermore, students' responses to the ninth statement in the pre-survey reveal that over 80% of the participants felt that they were willing to put in more effort into a task in order to get further knowledge, with 55.6% (15 participants) stating that they were almost always willing to put extra effort and 40.7% (11 participants) of the learners stating that they often did so. Following the intervention, participants' responses to the same statement indicated that almost all participants believed that they were willing to take on more work in order to gain further knowledge. 81.5% of the participants (22 participants) stated that they almost always did so, and 14.8% of the participants (four participants) stated that they were often willing to do extra in order to learn more. Lastly, the findings for the tenth statement reveal that, in the pre-survey, the majority of the participants (25 participants) believed that they would eventually succeed if they continued to work hard on a task, with 66.7% of the participants (18 participants) stating that they almost always felt this way and 25.9% of the participants (seven participants) stating that they often felt this way. Notably, the number of participants who felt this way increased after the intervention and the responses in the post-survey revealed that all the participants believed that, if they persisted on a task, they

would eventually succeed, with 81.5% (22 participants) stating that they almost always felt this way, and 18.5% (five participants) of the learners stating that they often felt that they would succeed if they continued in completing a task.

Furthermore, students' responses to all 10 statements, which aimed to explore their perceptions of their efforts, are merged and presented as one single indicator to represent learners' perceptions of their effort prior to and after the intervention. Table 5-14 presents the descriptive statistics on the effort section as a whole from both the pre- and post-survey, including the mean of the full sample (n=27) for each statement.

Table 5-14 Descriptive statistics of the full sample scores of the effort section in the pre- and post SRL-SRS survey

	N	Pre-survey					Post-survey				
		Mean	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
Effort	27	3.13	3.40	.56	1.80	4.00	3.53	3.60	.35	2.40	3.90

Table 5-14 that the mean of participants' responses in the post-SRL-SRS survey is higher than their responses in the pre survey, and the mean difference is 0.39. Nonetheless, in order to investigate whether the difference in means between the pre and post is statistically significant or not, the Wilcoxon Signed Rank Test and the findings are presented in Table 5-15.

Table 5-15 Statistical difference of the effort section in the pre and post-SRL-SRS survey

Test Statistics	
	Post-effort – Pre-effort
Z	-3.94[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-15 presents the results of Wilcoxon's Signed Rank Test, which indicate a statistically significant difference between learners' perceptions of their effort processes prior to and after the intervention. The test reveals a statistically significantly positive change in learners' perceptions, $Z = -3.94$, $p < .001$, with a large effect size ($r = .75$). The median score on the perceptions of effort increased from pre-survey ($Md = 3.40$) to post-survey ($Md = 3.60$).

These findings may be considered an indication of the positive impact of self-assessment and the intervention on learners' effort. The findings reveal positive changes in learners' perception of the level of dedicated effort that they put into their work after the intervention. One possible justification for these findings is that learners' active engagement and involvement in their speaking classes encouraged them to take a more active role in their learning and, thus, enhance the level of effort they put in their work. Nonetheless, an in-depth exploration of learners' perceptions through interviews can provide reasons for these results.

The findings and descriptive statistics of the sixth section in the survey are presented and discussed in the following sub-section.

5.2.6. EFL learners' perceptions of their self-efficacy in the SRL-SRS survey

This section reports findings on the sixth and final section of the SRL-SRS survey. The sixth section in the survey includes 10 statements that aimed to explore EFL learners' perceptions of the level of their self-efficacy. The learners indicated their views on a five-point scale; the higher the scale, the higher the confidence and high self-efficacy. Table 5-16 presents the descriptive statistics from both the pre- and post-survey, including the frequency and percentages of the full sample (n=27) for each statement.

Table 5-16 Descriptive statistics of the full sample's perceptions of their effort in the pre- and post-SRL-SRS survey

Statements	N		Pre-survey					Post-survey				
			Almost never	sometimes	often	Almost Always	Total	Almost never	sometimes	often	Almost Always	Total
1. I know how to handle unforeseen situations, because I can well think of strategies to cope with things that are new to me.	27	F	3	11	9	4	27	0	3	13	11	27
		%	11.1%	40.7%	33.3%	14.8%	100%	0%	11.1%	48.1%	40.7%	100%
2. If someone opposes me, I can find means and ways to get what I want.	27	F	2	9	8	8	27	0	5	8	14	27
		%	7.4%	33.3%	29.6%	29.6%	100%	0%	18.5%	29.6%	51.9%	100%
3. I am confident that I could deal efficiently with unexpected events.	27	F	2	12	7	6	27	0	2	15	10	27
		%	7.4%	44.4%	25.9%	22.2%	100%	0%	7.4%	55.6%	37.0%	100%
4. If I am in a bind, I can usually think of something to do.	27	F	2	5	14	6	27	0	3	11	13	27
		%	7.4%	18.5%	51.9%	22.2%	100%	0%	11.1%	40.7%	48.1%	100%
5. I remain calm when facing difficulties, because I know may ways to cope with difficulties.	27	F	2	12	11	2	27	0	4	15	8	27
		%	7.4%	44.4%	40.7%	7.4%	100%	0%	14.8%	55.6%	29.6%	100%
6. I always manage to solve difficult problems if I try hard enough.	27	F	0	4	13	10	27	0	1	7	19	27
		%	0%	14.8%	48.1%	37.0%	100%	0%	3.7%	25.9%	70.4%	100%
7. I can solve most problems if I invest the necessary effort.	27	F	0	11	6	10	27	0	1	7	19	27
		%	0%	40.7%	22.2%	37.0%	100%	0%	3.7%	25.9%	70.4%	100%
8. It is easy for me to concentrate on my goals and to accomplish them.	27	F	2	11	7	7	27	0	4	10	13	27
		%	7.4%	40.7%	25.9%	25.9%	100%	0%	14.8%	37.0%	48.1%	100%
9. When I am confronted with a problem, I usually find several solutions.	27	F	2	14	9	2	27	0	1	19	7	27
		%	7.4%	51.9%	33.3%	7.4%	100%	0%	3.7%	70.4%	25.9%	100%
10. No matter what comes my way, I'm usually able to handle it.	27	F	0	5	16	6	27	0	2	11	14	27
		%	0%	18.5%	59.3%	22.2%	100%	0%	7.4%	40.7%	51.9%	100%

The data in Table 5-16 reveals changes in learners' perceptions of their self-efficacy prior to and after the intervention. The data presented in Table 5-16 show learners' responses to first statements, and reveal the number of participants who believed that they were capable of dealing with unanticipated circumstances since they were capable of finding solutions to deal with new challenges after the intervention. In the post-survey, 88.8% of the participants (24 participants) thought that they knew how to handle unforeseen situations because they had strategies to cope with things that were new to them, with 40.78% of the participants (11 participants) stating almost always and 48.1% (13 participants) stating that they often felt this way in comparison to 48.1% (13 participants) in the pre-survey with 14.8% (four participants) almost always and 33.3% (nine participants) often feeling that they were capable of dealing with unanticipated circumstances and finding ways for coping with unfamiliar situations. Moreover, learners' responses to the second statement in the pre-survey reveal that over 59.2% of the participants believed that, if someone challenged them, they would be capable of finding means and methods to get what they wanted, with 29.6% (eight participants) stating that they were often able to find the means to get what they want and 29.6% (eight participants) of the learners stated that they almost always did so. The participants' responses to the same statement after the intervention reveal that over 80% of the participants felt that, if they were challenged, they would be capable of devising tactics to get what they wanted. 51.9% of the participants (14 participants) stated that they almost always did so and 29.6% of the participants (eight participants) state that they often felt so. Table 5-16 also presents learners' responses to the third statement in the pre-survey and post-survey, revealing that, after the intervention, more learners had confidence that they could cope with unforeseen circumstances effectively. Table 5-16 reveals that, in the post-survey, 78.6% of the participants (25 participants) were confident in their ability to cope with unforeseen situations effectively, with 23.0% of the participants (10 participants) stating almost always and 55.6% (15 participants) stating that they often felt this way in comparison to 48.1% (13 participants) in the pre-survey, with 22.2% (six participants) stating almost always and 25.9% (seven participants) often feeling confident that they could deal efficiently with unexpected events.

Furthermore, Table 5-16 shows students' responses to the fourth statement in the pre-survey, revealing that 74.1% of the participants believed that, when in a dilemma, they could usually think of something to do, with 51.9% (14 participants) stating that they often had an idea, and 22.2% (six participants) of the learners stating that they almost always did so. The participants' responses to the same statement in the post-survey reveal that 88.8%

of the participants believed that, when in a bind, they could come up with something to do. 48.1% of the participants (13 participants) stated that they almost always did so, and 40.7% of the participants (11 participants) stated that they often did so. The data presented in Table 5-16 also show learners' responses to the fifth statements, revealing that the number of participants who felt that they remained calm when confronted with obstacles, since they are aware of several strategies for dealing with them increased after the intervention. In the post-survey, a total of 85.2% of the participants (23 participants) felt that, because they know many ways to cope with difficulties, remain calm when facing difficulties, with 29.6% of the participants (eight participants) stating almost always and 55.6% (15 participants) stating that they often felt this way in comparison to 48.1% (13 participants) in the pre-survey, with 7.4% (two participants) stating almost always and 40.7% (11 participants) often feeling that they knew many ways to cope with difficulties.

Table 5-16 presents learners' responses to the sixth statement, which reveal that in the pre-survey, 85.1% of the participants believed that they usually succeeded in solving challenging problems if they worked hard enough, with 37.0%% (10 participants) almost always and 48.1% (13 participants) often feeling so. Notably, in the post-survey, the majority of the participants (26 participants) felt that, if they worked hard enough, they always succeeded in solving challenging problems, with 70.4% of the participants (19 participants) stating almost always and 25.9% (seven participants) often feeling this way. Furthermore, the table presented learners' responses to the seventh statement, revealing that the number of participants who believed that they were capable of solving most issues if they made appropriate effort increased after the intervention. In the post-survey, the majority of the participants (26 participants) believed that, if they made the required effort, they were capable of resolving most problems, with 70.4% of the participants (19 participants) stating almost always and 25.9% (seven participants) often feeling this way in comparison to 59.2 (16 participants) in the pre-survey, with 37.0% (10 participants) almost always and 22.2% (six participants) often feeling that they could solve most problems if they invested the necessary effort.

The data presented in Table 5-16 also reveal learners' responses to the eighth statement in the pre-survey. 51.8% of the participants believed that concentrating on and achieving their objectives was effortless to them, with 25.9% (seven participants) stating that, when concentrating on their objectives, achieving these objectives became effortless. Moreover, 25.9% (seven participants) of the learners stated that they almost always felt this way. In the post-survey, 85.1% of the participants believed that it was easy for them to concentrate

on their goals and to accomplish them, with 48.1% of the participants (13 participants) stating that they almost always felt so and 37.0% of the participants (10 participants) stating that they often felt so.

Furthermore, Table 5-16 presents learners' responses to the ninth statement. In the pre-survey, 40.7% of the participants believed that, when they are confronted with a problem, they usually found several solutions, with 7.4% (two participants) stating almost always and 33.3% (nine participants) stating that they often felt so. Remarkably, in the post-survey, the majority of the participants (26 participants) felt that they usually found several solutions when they faced a problem, with 70.4% of the participants (19 participants) stating often and 25.9% (seven participants) almost always feeling this way. Lastly, learners' responses to the tenth statement reveal that the number of participants who believed that they were capable of dealing with any issue that they faced increased after the intervention. In the post-survey, 92.6% of the participants (25 participants) felt that, no matter what came their way, they were usually able to handle it, with 51.9% of the participants (14 participants) stating almost always and 40.7% (11 participants) stating that they often felt this way in comparison to 81.5% (22 participants) in the pre-survey with 22.2% (six participants) stating almost always and 59.3% (16 participants) often feeling that they knew how to handle any issue.

The descriptive statistics of learners' responses to all 10 statements presented as one single indicator to represent learners' perceptions of their self-efficacy, prior to and after the intervention, are presented in Table 5.17.

Table 5-17 Descriptive statistics of the full sample scores of the self-efficacy section in the pre- and post SRL-SRS survey

	N	Pre-survey					Post-survey				
		Mean	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
Self-efficacy	27	2.77	2.80	.48	1.50	3.70	3.37	3.40	.33	2.60	4.00

Based on, it is evident that the mean of participants' responses in the self-efficacy section in the post-survey is higher than their responses to the pre-survey, and the mean difference is 0.61. However, as discussed in the previous subsection, Wilcoxon's Signed Rank Test was used to investigate whether the difference in means between learners' perceptions of their self-efficacy in the pre- and post- survey is statistically significant. These findings are presented in Table 5-18.

Table 5-18 Statistical difference of the self-efficacy section in the pre and post-SRL-SRS survey

Test Statistics	
	Post- self-efficacy – Pre- self-efficacy
Z	-4.50[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-18 presents the results of Wilcoxon's Signed Rank Test to reveal a statistically significant difference between learners' perceptions of their self-efficacy processes prior to and after the intervention. The test found a statistically significantly positive change in learners' perceptions, Z= -4.50, $p < .001$, with a large effect size (r = .86). The median score on the perceptions of self-efficacy increased from pre-survey (*Md* = 2.80) to post-survey (*Md*= 3.40).

By considering these statements on self-efficacy, learners' perceptions of the level of their self-efficacy increased after the intervention. Learners' responses reveal a high level of confidence and independency. It is impossible to be certain of reasons for this positive change in learners' attitude and perceptions before exploring their views and experience in depth through interviews. One possible explanation for these changes is that, through self-assessment, learners were able to be more involved in their learning and saw results of improvement, and believed in their ability to overcome difficulties, which raised their self-efficacy.

The findings and descriptive statistics of the overall perceptions of self-regulated learning including all six sections in the survey, are presented and discussed in the following sub-section.

5.2.7. EFL learners' overall perceptions of their self-regulated learning in the SRL-SRS survey

This section reports learners' overall perceptions of all six sections in the SRL-SRS survey. The descriptive statistics of learners' responses to all 10 statements presented as one single indicator to represent learners' perceptions of their self-efficacy, prior to and after the intervention are presented in Table 5-19.

Table 5-19 Descriptive statistics of the full sample scores of the overall SRL from the pre- and post SRL-SRS survey

	N	Pre-survey					Post-survey				
		Mean	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**
Overall SRL	27	3.29	3.37	.38	2.16	3.87	3.91	3.94	.25	3.22	4.25

Table 5-19 shows that the mean of the overall responses in the post-SRL-SRS survey is higher than the overall responses in the pre survey, and the mean difference is 0.62. Nonetheless, in order to investigate whether the difference in means between the pre and post is statistically significant or not, the Wilcoxon Signed Rank Test and the findings are presented in Table 5-20.

Table 5-20 Statistical difference of the overall SRL in the pre and post-SRL-SRS survey

Test Statistics	
	Post- overall SRL – Pre- overall SRL
Z	-4.54[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-20 presents the results of Wilcoxon's Signed Rank test, which indicate a statistically significant difference between learners' perceptions of the overall processes prior to and after the intervention. The test reveals a statistically significantly positive change in learners' perceptions, $Z = -4.54$, $p < .001$, with a large effect size ($r = .87$). The median score on the perceptions of overall SRL increased from pre-survey ($Md = 3.36$) to post-survey ($Md = 3.93$).

In summary, the results of the SRL-SRS prior to and after the intervention reveal positive changes in learners' attitudes toward their self-regulated learning skills. These results could be considered an indication that self-assessment has a positive impact of on EFL learners' self-regulated learning skills, though it is not possible to determine the positive impact of self-assessment before exploring and understanding learners' perceptions and experience in-depth through interviews.

Moreover, based on the results of the SRL-SRS, evaluation had the largest effect size between the six self-regulated learning skill included in SRL-SRS (r = .62). This finding indicates that self-assessment has a higher positive impact on learners' views about their evaluation skills than other skills.

The following section presents and discusses the finding of the critical thinking test before and after the intervention.

5.3. EFL learners' critical thinking skills

As discussed in Chapter 4, the Watson-Glaser II Critical Thinking Appraisal (CTA) was used in this study to measure participants' critical thinking before and after the intervention, and the total number of respondents was n=27. Data management and analysis were performed using SPSS software (Version 26). As the first stage of the analysis, the pre- and post-CTAs of the full sample of participants were corrected using the answer sheets provided with the test by TalentLens, the providers of the test. The grade for each section was calculated along with the total for all five sections for both the pre- and post-CTAs.

Moreover, before describing the test used for data analysis, it should be noted that the CTA is measured through grades described as continuous variables as a ratio scale of measurement (Pallant, 2010; Pallant, 2016; Salkind, 2010). The participants' grades for all five sections, along with the total grade of the CTA were first analysed using descriptive statistical methods. The means, medians, and standard deviation were used to provide descriptive statistics for continuous variables (Pallant, 2010; Pallant, 2016; Salkind, 2010). Data were then analysed using inferential statistics as a mean to determine if there were statistically significant difference. Parametric tests are often used to analyse continuous variables; however, since the data did not meet the assumption of normality required for parametric test, non-parametric tests were applied to perform the analysis of the CTA before and after intervention. Therefore, Wilcoxon's Signed Rank Test was used in this study to determine whether or not there were statistically significant differences between two related samples, namely learners' critical thinking in the pre- and post-surveys. Both descriptive and inferential statistics of the CTA prior to and after the intervention are discussed.

The data presented in Table 5-21 Descriptive statistics of the full sample's grades of the infer section in the pre- and post-CTA show learners' grades for the first section of the CTA, which aimed to assess learners' ability to make inferences by having them determine the degrees of truth or untruth of an inference. The results reveal an increase of learners' grades in this section after the intervention. In details, the data reveal that learners' grades increased from the pre-CTA (M =2.26, SD = .98) to post-CTA (M =3.44, SD = 1.19).

Table 5-21 Descriptive statistics of the full sample's grades of the infer section in the pre- and post-CTA

	N	Pre-CTA					Post-CTA				
		Min	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**
Infer	27	1	5	2.26	2.00	.98	1	5	3.44	4.00	1.19

However, the significance of the difference cannot be ascertained only with descriptive statistics. Therefore, Wilcoxon's Signed Rank Test was used to determine any statistically significant difference between learners' grades in this section in the pre- and post-CTA, the findings are presented in Table 5-22.

Table 5-22 Statistical difference of the infer section in pre- and post-CTA

Test Statistics	
	Post-infer – Pre-infer
Z	-3.99[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-22 presents the results of Wilcoxon's Signed Rank Test presented, which reveal a statistically significant difference between learners' grades prior to and after the intervention. The test indicates a statistically significantly positive improvement in learners' grades, Z= -3.99, $p < .001$, with a large effect size (r = .76). The median score of the learners' grades in this section increased from pre-CTA (*Md* = 2.00) to post-CTA (*Md*= 4.00).

Moreover, Table 5-23 presents learners' grades for the second section of the CTA, which aimed to assess their ability to recognise underlying assumptions or presuppositions. The findings reveal an increase in learners' grades after the intervention. Furthermore, the data reveal that learners' grades increased from the pre-CTA (M =6.15, SD = .2.01) to post-CTA (M =8.30, SD = 2.03).

Table 5-23 Descriptive statistics of the full sample's grades of the recognise assumptions section in the pre- and post-CTA

	N	Pre-CTA					Post-CTA				
		Min	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**
Recognise Assumption	27	2	9	6.15	7.00	2.01	4	11	8.30	9.00	2.03

Moreover, in order to determine a statistically significant difference between learners' grades in this section in the pre- and post-CTA, Wilcoxon's Signed Rank Test was used, the findings of which are presented in Table 5-24

Table 5-24 Statistical difference of the recognize assumptions section in pre- and post-CTA

Test Statistics	
	Post- Recognise Assumptions – Pre- Recognise Assumptions
Z	-3.68[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-24 presents the results of Wilcoxon's Signed Rank test to reveal statistically significant differences between learners' grades prior to and after the intervention. The test indicates a statistically significantly positive improvement in learners' grades, Z= -3.68, $p < .001$, with a large effect size ($r = .70$). The median score of the learners' grades in this section increased from pre-CTA (*Md* = 7.00) to post-CTA (*Md*= 9.00).

Learners' grades in the third section are presented in Table 5-25. This section aimed to assess learners' deductive ability by asking them to determine whether or not specific facts were definitely true. The findings presented in Table 5-25 reveal an increase in learners' grades after the intervention. In details, learners' grades increased from the pre-CTA (M =3.04, SD = 1.16) to post-CTA (M =4.00, SD = 1.04).

Table 5-25 Descriptive statistics of the full sample's grades on the deduce section in the pre- and post-CTA

	N	Pre-CTA					Post-CTA				
		Min	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**
Deduce	27	0	5	3.04	3.00	1.16	1	5	4.00	4.00	1.03

Wilcoxon's Signed Rank Test was used to determine any statistically significant difference between learners' grades in this section in the pre- and post-CTA. These findings are presented in Table 5-26.

Table 5-26 Statistical difference of the deduce section in pre- and post-CTA

Test Statistics	
	Post-Deduce – Pre-Deduce
Z	-3.36[b]
Asymp. Sig. (2-tailed)	.001
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-26 presents the results of Wilcoxon's Signed Rank Test to reveal a statistically significant difference between learners' grades prior to and after the intervention. The test found a statistically significantly positive improvement in learners' grades, Z= -3.36, p = .001, with a large effect size (r = .64). The median score of the learners' grades in this section increased from pre-CTA (Md = 3.00) to post-CTA (Md= 4.00).

Moreover, Table 5-27 outlines learners' grades for the CTA's fourth section, which aimed to assess interpretive ability by having students evaluate evidence and determine if generalisations or conclusions were appropriate. The data presented in Table 5-27 reveal an increase in learners' grades after the intervention. In details, the data revealed that learners' grades improved from the pre-CTA (M =2.70, SD = 1.17) to post-CTA (M =3.96, SD = 1.05).

Table 5-27 Descriptive statistics of the full sample's grades of the interpret section in the pre- and post-CTA

	N	Pre-CTA					Post-CTA				
		Min	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**
Interpret	27	0	5	2.70	3.00	1.17	1	6	3.96	4.00	1.05

However, Wilcoxon's Signed Rank Test was used to determine any statistically significant difference between learners' grades in this section in the pre- and post-CTA, the findings of which are presented in Table 5-28.

Table 5-28 Statistical difference of the interpret section in pre- and post-CTA

Test Statistics	
	Post- Interpret – Pre-Interpret
Z	-3.72[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-28 presents the results of Wilcoxon's Signed Rank Test to reveal a statistically significant difference between learners' grades prior to and after the intervention. The test found a statistically significantly positive improvement in learners' grades, $Z = -3.72$, $p < .001$, with a large effect size ($r = .71$). The median score of the learners' grades in this section increased from pre-CTA (*Md* = 3.00) to post-CTA (*Md*= 4.00).

Furthermore, Table 5-29 presents learners' grades in the fifth section in the CTA prior to and after the intervention, which aimed to determine if students were capable of differentiating between strong and relevant arguments and weak or irrelevant arguments. The data reveal an increase of learners' grades after the intervention from the pre-CTA (M =6.37, SD = 1.24) to post-CTA (M =8.52, SD = 1.37).

Table 5-29 Descriptive statistics of the full sample grades of the evaluation argument section in the pre and post CTA

	N	Pre-CTA					Post-CTA				
		Min	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**
Evaluation Arguments	27	4	10	6.37	6.00	1.24	5	11	8.52	9.00	1.37

Wilcoxon's Signed Rank Test was then used to determine a statistically significant difference between learners' grades in this section in the pre- and post-CTA, the findings of which are presented in Table 5-30.

Table 5-30 Statistical difference of the evaluation arguments section in pre- and post-CTA

Test Statistics	
	Post- Evaluation Arguments – Pre- Evaluation Arguments
Z	-4.01[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-30 presents the results of Wilcoxon's Signed Rank Test to reveal a statistically significant difference between learners' grades prior to and after the intervention. The test found a statistically significantly positive improvement in learners' grades, Z= -4.01, $p < .001$, with a large effect size (r = .77). The median score of the learners' grades in this section increased from pre-CTA (*Md* = 6.00) to post-CTA (*Md*= 9.00).

Furthermore, the descriptive statistics of the total test score of the five sections in the pre- and post- CTA are presented in Table 5-31. Based on Table 5-31, there was an increase in learners' grades after the intervention. In details, the data reveal that learners' total grades improved from the pre-CTA (M =20.52, SD = 3.99) to post-CTA (M =28.44, SD = 4.59).

Table 5-31 Descriptive statistics of the full sample grades of the total score in the pre and post CTA

	N	Pre-CTA					Post-CTA				
		Min	**Max**	**Mean**	**Median**	**SD**	**Min**	**Max**	**Mean**	**Median**	**SD**
CT Total score	27	10	28	20.5	22.0	3.99	19	36	28.4	30.0	4.59

Moreover, in order to determine any statistically significant difference between learners' grades in this section in the pre- and post-CTA, the Wilcoxon Signed Rank Test was used, these findings of which are presented in Table 5-32.

Table 5-32 Statistical difference of the total score in pre- and post-CTA

Test Statistics	
	Post- total score – Pre- total score
Z	-4.31[b]
Asymp. Sig. (2-tailed)	.000
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 5-32 presents the results of the Wilcoxon Signed Rank Test presented and it reveals a statistically significant difference between learners' grades prior to and after the intervention. The test found statistically significantly positive improvement in learners' grades, $Z= -4.31$, $p < .001$, with a large effect size ($r = .82$). The median score of the learners' grades in this section increased from pre-CTA (Md = 22.0) to post-survey (Md= 30.0).

Moreover, based on the results of the CTA, the inference and evaluation arguments critical thinking skills displayed the highest effect size between the five critical thinking skills with a large effect size ($r = .77$). The outcome of both CTA, prior to and after the intervention, showed an increase in learners' grades on the test measuring their critical thinking abilities. One possible explanation for these changes is the impact of self-assessment on learners' critical thinking abilities, which is discussed further in Chapter 7. Another possible explanation is the relationship between learners' self-regulated learning skills and their critical thinking abilities. Therefore, the relationship between learners' perceptions of their self-regulated learning and critical thinking is examined and explored in the following section.

5.4. The relationship between EFL learners' perceptions of their self-regulated learning and their critical thinking skills

After determining a statistically significant change in learners' perceptions of their self-regulated learning skills and improvement in their critical thinking, it was necessary to determine the relationship between participants' self-regulated learning skills and critical thinking, as one of the main purposes of this research was to examine the extent of the relationship between these variables. According to Cohen et al. (2018), Spearman's rho is the non-parametric test used for investigating connections among variables. Spearman's rho values vary from -1.0 to +1.0, where the greater the connection between variables, the larger the coefficient (Cohen et al., 2018). When two variables fluctuate in the same direction, i.e. when one increases, the other increases, or vice versa, and a positive connection or correlation occurs. Therefore, +1.0 represents a perfect positive correlation variable (Cohen et al., 2018). In contrast, a negative connection or correlation occurs when one variable increases while the other decreases. Therefore, -1.0 represents a perfect negative correlation between variables (Cohen et al., 2018). However, perfect correlations are uncommon in social research, with the majority of correlation coefficients falling at approximately +0.5 or less (Cohen et al., 2018). Finally, a connection between variables

with a value of zero implies that there is no relationship (Cohen et al., 2018). Spearman's rho was used to explore whether there was a relationship between EFL learners' self-regulated learning skills and critical thinking for the full sample (n=27) and to explore whether there was interdependency between self-regulated learning skills. This test was performed at two points of this study, namely before and after the intervention, the results of which are presented in Table 5-33.

Table 5-33 Spearman's rho correlation matrix between self-regulated learning skills and critical thinking of the full sample prior to and after the intervention

			Pre-intervention							Post-intervention							
	Variable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Spearman's rho	**1. CT**																
	2. SRL-1	-.16								.38*							
	3. SRL-2	-.07	.42*							.21	.75**						
	4. SRL-3	-.09	.49**	.69**						.26	.46*	.49**					
	5. SRL-4	.16	.14	-.01	.30					.5	.52**	.35	.49**				
	6. SRL-5	-.29	.49**	.69*	.78**	.18				.17	.47*	.62**	.57**	.14			
	7. SRL-6	-.05	.53**	.52**	.49**	.12	.57**			.38	.49**	.45*	.31	-.00	.60**		
	8. SRL-7	-.09	.71**	.69**	.83**	.44**	.83**	**.**71**		.36	.85**	.81**	.73**	.55**	.71**	**.**68**	

(**CT**: Critical thinking, **SRL-1**: Self-regulated learning- Planning, **SRL-2**: Self-regulated learning- Self-monitoring, **SRL-3**: Self-regulated learning- Evaluation, **SRL-4**: Self-regulated learning- Reflection, **SRL-5**: Self-regulated learning- Effort, **SRL-6**: Self-regulated learning- Self-Efficacy, **SRL-7**: SRL- overall).

*correlation is significant at the 0.05 (2-tailed)

**correlation is significant at the 0.01 (2-tailed)

The results of Spearman's rho for the CT and SRL skills before and after the intervention are presented in Table 5-33. Prior to the intervention, there were weak negative relationships between learners' critical thinking and learners' self-regulated learning skills of planning (SRL-1), self-monitoring (SRL-2), evaluation (SRL-3), effort (SRL-5), and self-efficacy (SRL-6), except for reflection (SRL-4), which had a weak positive relationship with learners' CT. However, the data presented in Table 5-33 reveal that the overall relationship between CT and SRL skills is a weak positive relationship, and is not statically significant.

This relationship changed after the intervention, indicating a medium-strength positive relationships between CT and SRL. Also, the relationship between CT and SRL skills shifted after the intervention, indicating weak positive relationships between CT and self-monitoring (SRL-2), evaluation (SRL-3), and effort (SRL-5) and medium-strength positive relationships between CT and planning (SRL-1), reflection (SRL-4), and self-efficacy (SRL-6). Additionally, Table 5-33 revealed that the relationship between critical thinking and planning skill is statistically significant, $p < .05$.

Moreover, Table 5-33 presents relationships before the intervention between the self-regulated learning skills that vary between weak, medium strength, and strong positive relationships and one weak negative relationship between reflection (SRL-4) and evaluation. Additionally, most of these positive relationships are statistically significant, $p < .05$ or $p < .01$. Notably; however, the strength of the relationship between SRL skills increased after intervention. The table suggests that the majority of the skills have a medium or positive relationship. The table also reveals a weak relationship between effort (SRL-5) and reflection (SRL-4), as well as a weak negative relationship between effort (SRL-5) and self-efficacy (SRL-6).

5.5. Conclusion

The findings presented in this chapter were gathered by two pre- and post- tests, SRL-SRS and CTA, that explored EFL students' perspectives of their self-regulated learning practices and measured their critical thinking before to and after self-assessment intervention. The quantitative data were analysed using descriptive statistics and a non-parametric test, namely Wilcoxon Signed Rank. Overall, the findings from the SRL-SRS reveal favourable shifts in the learners' perceptions of their self-regulated learning ability after the self-assessment intervention. These findings might be supported by the findings

from the in-depth investigation of the learners' perspectives through interviews, which will be analysed in Chapter 6. Similarly, the overall findings from the CTA showed improvement in students' results in terms of their critical thinking skills.

Lastly, the relationship between self-regulated learning skills and critical thinking skills were investigated using Spearman's r correlation test, which revealed a strong positive medium-strength relationship between the two concepts.

The findings from the pre- and post-test are discussed in more detail in Chapter 7 in relation to the literature, with a view to answering the research questions.

Chapter 6 : Quantitative and Qualitative Data Analysis

6.1. Introduction

Chapter 4 outlined the research methodology for this study, including the use of mixed methods to explore and understand the EFL learners' lived experiences and perspectives of self-assessment and its impact on their self-regulated learning, critical thinking, and achievement in English speaking classrooms at the Saudi higher education institution (HEI) through pre and post surveys, first and last audio recordings, self-assessment proformas, and semi-structured interviews. This chapter begins by outlining the method and process used in the analysis of learners' audio recordings and self-assessment proformas. The chapter then discusses the method and process used in the analysis of the semi-structured interviews with the EFL learners. As discussed in Chapter 4 Section 4.6.2.4, the purpose of the interviews was to identify and explore the learners' experience with the use of self-assessment and the learners' perceptions of the use and impact of self-assessment in an English speaking classroom. The chapter then presents the themes that emerged from the interviews and interprets the findings.

6.2. Analysis Method and Process

This study aimed, in part, to explore EFL learners' perceptions of their lived experience with self-assessment and the impact of self-assessment on their self-regulated learning, critical thinking, and achievement through interviews. First, however, it was important to identify and explore whether the EFL learners were actively engaging in self-assessment. In addition, it had to be determined whether the learners used SRL strategies while completing the task. Finally, it also had to be determined whether the learners had improved in their performance and achievement. All of these factors provided valuable insight into learners' perceptions. The analysis and the results are presented below.

6.2.1. Audio recordings

Audio recordings of learners' first and last speaking task were collected during the intervention in order to uncover whether a significant statistical difference could be detected in learners' level of achievement after the implementation of self-assessment. As a first step of analysis, each audio recording was listened to multiple times and was graded twice at different times using the institution speaking rubric. Then, the average of the two

sets of data and the actual grade given by the teacher was calculated. The descriptive statistics of the participants' grades for both the first and last audio recordings are presented in Table 6-1.

Table 6-1 Descriptive statistics of speaking achievement of 10 EFL learners

	N	Pre-CTA					Post-CTA				
		Min	Max	Mean	Median	SD	Min	Max	Mean	Median	SD
Achievement	10	2.25	4.58	3.60	3.62	.61	3.50	4.92	4.31	4.33	.50

It is evident from Table 6-1 that the mean of participants' grades in the last audio recording is higher than the mean of their grades in the first recording. In order to confirm whether the difference in means is statistically significant, the Wilcoxon Signed Rank Test was used, the findings of which are presented in Table 6-2.

Table 6-2 Statistical difference of achievement in the first and last audio recordings

Test Statistics	
	1st recording – last recording
Z	-2.81b
Asymp. Sig. (2-tailed)	.005
a. Wilcoxon Signed Ranks Test b. Based on negative ranks.	

Table 6-2, which presents the results of the Wilcoxon Signed Rank Test, indicates a statistically significant difference between the participants' speaking grades prior to and after the intervention. The test revealed a statistically significant positive change in learners' grades, $Z = -2.81$, $p < .01$, with a large effect size ($r = .88$). The median score of the grades increased from the first recording ($Md = 3.62$) to the last recording ($Md = 4.33$).

6.2.2. Self-assessment proforma

Self-assessment proformas was collected to explore the EFL learners' usage of SRL strategies while practising self-assessment. The proformas were thus employed to capture traces of what learners really did, rather than just what learners claim to have done (Perry, 1998; Winne & Jamieson-Noel, 2002). Table 6-3 indicates whether there was any evidence of SRL strategies being employed; the table includes the SRL strategies of planning, self-monitoring, reflection, and evaluating.

Table 6-3 Samples of traces of the use of SRL Strategies from self-assessment proforma

Participants / SRL skills	S1								S11								S21							
	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
Planning	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Monitoring	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Evaluating	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Reflection	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	

Overall, the data revealed that all learners actively engaged with the intervention and employed all four SRL skills almost every time they used the self-assessment proforma, with a few minor incidents in which learners did not use the reflection skill.

The following section addresses the methods and processes of analysing the interviews.

6.2.3. Interviews Analysis Method and Process

The semi-structured interview data in this research was analysed using thematic analysis. Braun and Clarke (2006) referred to thematic analysis as the essential foundation of qualitative analysis. Thematic analysis is generally regarded as a key technique for qualitative analysis since it provides researchers with a range of core abilities for conducting any type of qualitative analysis. Thematic analysis is adaptable, and as such, it is a useful tool for analysing categories in order to create themes (Alhojailan, 2012) and provide comprehensive interpretation of data (Braun & Clarke, 2006). Moreover, thematic analysis may be adapted or organised to align with the interpretivist approach, which aims to convey participants' experiences, ideas, and realities (Braun & Clarke, 2006). Thus, thematic analysis is applicable to any research that aims at exploration and interpretation since it includes a systematic approach for data analysis (Alhojailan, 2012) and enables the researcher to understand any topic or issue (Marks & Yardley, 2004).

In thematic analysis, a theme can be defined in terms of what may be important to the research questions; a theme may not necessarily be easily quantified. Furthermore, many researchers have indicated that themes and patterns in thematic analysis may be defined

inductively (bottom up), wherein codes are derived directly from data (data-driven), or deductively (top down), wherein codes are set prior and stem from the research aim and questions and literature (theory-driven) (Boyatzis, 1998; Braun & Clarke, 2006; Hayes, 1997, 2013). For this research, the method of thematic analysis selected was inductive interpretative phenomenological analysis (IPA). Specifically, the research applied the IPA stages for thematic analysis as described by Smith and Osborn (2003), Eatough and Smith (2017), and Smith and Shinebourne (2012).

6.2.3.1. Interpretative Phenomenological Analysis

IPA, as per the stages of the aforementioned researchers, follows three steps.

Step 1: Search for initial themes in all cases.

The first step of IPA requires the researcher to become familiar with the data through focused repeated reading of the transcript, as each reading has the potential of eliciting new perceptions (Smith & Osborn, 2003). In the present research, the transcripts were printed, and the analysis was conducted manually on paper; as suggested by Smith and Osborn (2003), the left-hand margin of the transcript was used to comment and annotate what was noteworthy or important from each participant's statements. Remarks in the transcripts had several purposes. Some represented efforts to summarise or paraphrase. Others noted associations or links that came to mind, attempted early interpretations, or highlighted participants' language use, or the sense of the participants that is coming across. This process, which was undertaken for all interview transcripts, gave me a preliminary understanding of the overall perspectives and outcomes. Next, the right-hand margin was used to write emerging theme titles. Themes emerged from the preliminary notes, which were transformed into initial themes or concise phrases with the purpose of capturing the essence of the outcomes and what was found in the transcript (Smith & Osborn, 2003). Writing these initial themes and phrases transitioned the analysis to a slightly higher level of abstraction; namely, though I was still grounded in the particularity of the specific words said, it was necessary to find expressions high level enough to capture theoretical connections within and across cases (Smith & Osborn, 2003). This process of transforming preliminary notes into initial themes was conducted with all interview transcripts (See Appendix I for sample of the manual analysis of interviews). According to Smith and Osborn (2003), I treated all transcripts as data at this step, making no attempt to omit or select particular passages for special attention. In addition, not all preliminary notes turned

into or generated themes. The number of emerging themes in each transcript indicated the richness of the interview. The initial themes are shown in Table 6-4.

Table 6-4 Initial list of Themes

Initial list of Themes
Self-assessment requires practice.
Learners' active involvement in the learning process.
Positive feelings associated with self-assessment.
The development of metacognitive knowledge.
The development of cognitive processes.
The development of metacognitive strategies.
The development of critical thinking abilities.
The development of critical thinking dispositions.
Enhancing performance and achievement.
Relevance and appropriateness of assessment criteria
Informing and understanding speaking assessment requirements or criteria.
The purpose of assessment.
Speaking assessment as a de-motivator.
Anxiety associated with traditional speaking assessment.
The role of the teacher in the assessment process.
Support or against teacher feedback.
Support or against peer-assessment.
Support other sources of feedback.
Anxiety associated with receiving feedback.
The significance of feedback.
Learners' feedback expectations.

Step 2: Cluster themes.

The second step of IPA aims to identify clusters of themes that capture the participants' perceptions on a topic. I used sticky notes to explore and identify clusters of themes, writing the initial themes on sticky notes so that I could easily move them while searching for possible clusters. As noted by Smith and Osborn (2003), this step requires more analytical and theoretical analysis as the researcher attempts to make sense of the initial themes. In this step, some of the initial themes may merge into clusters, while others may develop as subordinate themes or cluster of themes (Smith & Osborn, 2003). For instance, the theme of "the benefits of self-assessment" was identified as an initial theme and then developed as a subordinate theme. Table 6-5 presents the clustering of themes. The clustered themes were then verified in the transcripts to ensure that the connections made sense according to the participants' actual words (Smith & Osborn, 2003). Additionally, according to Smith and Osborn (2003), the themes should be selected based on their occurrence and richness within the data. The clustered themes were checked across participants to verify that they are represented in more than half of the participants. The occurrences of subordinate themes are shown in Table 6-6. A sample of participants' quotes used to identify a theme are shown in Table 6-7.

Table 6-5 Clusters of initial themes

Cluster of themes	Initial themes
The use of self-assessment in speaking classes	Self-assessment requires training.
	Learners' active involvement in the learning process.
	Positive feelings associated with self-assessment.
The benefits of self-assessment	The development of metacognitive knowledge.
	The development of cognitive processes.
	The development of self-regulatory strategies.
	The development of critical thinking abilities and disposition.
	Enhancing performance and achievement.
The significance of the self-assessment proforma	Relevance and appropriateness of assessment criteria
	Informing and understanding speaking assessment requirements or criteria.
The use of traditional speaking assessments (exams)	The purpose of assessment.
	Speaking assessment as a de-motivator.
Challenges linked with traditional speaking assessment (exams)	Anxiety associated with speaking assessment.
	The role of the teacher in the assessment process.
The source of feedback	Support or against teacher feedback.
	Support or against peer-assessment and feedback.
	Support other sources of feedback.
Perceptions of receiving feedback	Anxiety associated with receiving feedback.
	The significance of feedback.
	Learners' feedback expectations.

Table 6-6 Sample of occurrences of cluster of themes

Occurrences of cluster of themes	S1	S3	S11	S18	S21
The use of self-assessment in speaking classes	✓	✓	✓	✓	✓
The benefits of self-assessment	✓	✓	✓	✓	✓
The significance of the self-assessment proforma	✓	✓	✓	✓	✓
The use of traditional speaking assessments (exams)	✓	✓	✓	✓	✓

Table 6-7 Sample of participants' quotes used to identify a theme

Cluster of themes	**Quotes from participants that identify the theme**
The use of traditional speaking assessments	I was very scared from judgement, because the normal way of evaluation was never enjoyable (S3).
	I had to memorise my answers even if I didn't understand them (S3).
	They (speaking assessment) were like the scariest, slowest and most stressful times (S3).
	Speaking usually is more like a test, we make groups and do the speaking, so it's more like a prepared script that we need to memorise and say it to the teacher (S11).
	In speaking classes, you don't really care if you did a good job or not because in the end all that actually mattered were the grades from the speaking (S11).

Step 3: Identify final themes.

The third step of IPA aims to connect clusters of themes and find patterns to identify final themes (superordinate themes). This involved the use of sticky notes to easily move and organise clusters of themes to identify final themes. Table 6-8 presents the list of final themes.

Table 6-8 List of final themes

Final themes	Cluster of themes
Perceptions and experiences of traditional speaking assessment	The use of traditional speaking assessments
	Challenges linked with traditional speaking assessment
Learners' perceptions and experiences of self-assessment in speaking classes	The use of self-assessment in speaking classes
	The benefits of self-assessment
	The significance of the self-assessment proforma
Perceptions and experiences of feedback in speaking classes	Source of feedback
	Perceptions of receiving feedback

The following section presents the themes that emerged from the learners' interviews.

6.3. Findings

The interviews were mainly conducted to gather deep information regarding Saudi EFL learners' perspectives on the effect of self-assessment in English speaking classes and their experiences with self-assessment. From the learners' interviews, three themes emerged. The first theme is learners' perceptions and experiences of traditional speaking assessment. The second theme is learners' perceptions and experiences of self-assessment in speaking classes. The third theme is learners' perceptions and experiences of feedback in speaking classes.

6.3.1. Learners' perceptions and experiences of traditional speaking assessment

This theme concerns learners' perspectives on the nature of assessment practices in English speaking classes and the strong emphasis in those classes on summative purposes of assessment. The participants' opinions and attitudes are discussed and presented through the following two subordinate themes: the use of traditional speaking assessments and the challenges linked with traditional speaking assessments.

6.3.1.1. The use of traditional speaking assessments

The interviews indicated that assessment in the learners' speaking classes occurs mainly in the form of oral presentations or exams conducted throughout the term. All EFL learners believed that assessment in speaking classes is mainly carried out for summative purposes, namely for evaluating and measuring their progress and performance. The learners did not feel that assessment aims to support them. For example, one participant stated that "speaking usually is more like a test, we make groups and do the speaking, so it's more like a prepared script that we need to memorise and say it to the teacher" (S11). As another participant noted, "I had to memorise my answers even if I didn't understand them" (S3). These statements suggest that speaking assessments led learners to rely on a lower-order cognitive skill in Bloom's taxonomy (Anderson & Krathwohl, 2001), namely memorising, and emphasised only the results. The learners further expressed that this form of assessment pressured them to concentrate on grades instead of on actually improving their language; as a result, the assessments eventually limited their learning.

> In speaking classes, you don't really care if you did a good job or not because in the end all that actually mattered were the grades from the speaking (S11).

> The normal way of evaluation was never enjoyable. They were like the scariest, slowest and most stressful times. I had to memorise my answers even if I didn't understand them, because all what we ever cared about were the marks (S3).

> Most of the English classes that I have participated in and like specifically speaking classes, I always end up caring about my grades at that moment and I rarely come out of that class with benefits (S1).

These quotations reveal a shared belief among the learners that speaking assessment hinder and even prevent learning opportunities and improvement. The quotations also reveal that speaking exams and the focus on grades can demotivate learners and discourage learning. The learners' use of phrases such as "you don't really care if you did a good job" or "I just want to get it over with and know my mark" indicate that assessment with only summative purposes can ultimately undermine learning opportunities.

6.3.1.2. Challenges linked with traditional speaking assessment

In addition, the data suggested that learners considered the speaking assessments to be a source of fear and anxiety. Learners' perceptions of traditional speaking assessments were linked to words like "demanding, scariest, difficult, and stressful", with one participant stating "the usual speaking assessment is really nerve-racking, especially the last one" (S2). This fear and concern seems to be connected to the value placed on grades. Learners highly priority their grades and seem to be fearful of their grades dropping, which would influence their GPA, affect their ability to choose their desired majors in university, and ultimately impact their future. Participants stated, for example, the following:

> The normal way of evaluation was never enjoyable. They were like the scariest, slowest and most stressful times. I had to memorise my answers even if I didn't understand them, because all what we ever cared about were the marks so I can enter the medical school and achieve my dream of being a doctor (S3).

> I work so hard on my-self and I'm always stressed that it will not be enough to have the future that my parent wish for me that is the reason why I feel terrified of being graded in every speaking assessment (S18).

> I always prepare for every evaluation and spend the time from one class to the next week class overthinking, worried, and stress because I want to have full mark but even after everything I do I lose marks (S21).

The statements above exemplify the learners' shared perception of speaking assessments as a cause of fear and anxiety. This fear of assessment may influence learners' motivation to learn and achieve language development. This negative influence can be deduced from learners' voices, which often became strained and expressed tension when the subject of speaking assessments and grades came up.

6.3.2. Learners' perceptions and experiences of self-assessment in speaking classes

The EFL learners' views of their traditional speaking assessment experiences made clear that the learners were not satisfied with the traditional form of assessment. Rather, the learners expressed feelings of discomfort and worry about the assessment, which they

believed was mainly conducted for summative purposes. However, although the learners questioned the role of traditional speaking assessments with purely summative purposes, the learners also recognised the importance of self-assessment and its role in helping them to be active learners in the learning process. This theme concerns how the learners described and viewed their experience with self-assessment in speaking classes. The learners' views and attitudes are discussed and presented through the following three subordinate themes: the use of self-assessment in speaking classes, the benefits of self-assessment, and the significance of the self-assessment proforma.

6.3.2.1. The use of self-assessment in speaking classes

The analysis revealed that all 10 participants had positive attitudes and experiences regarding self-assessment. In contrast to their views regarding traditional speaking assessment, learners' perspectives of self-assessment were linked to words like "comfortable, confident, liked, happy, interested, and encouraged". The following statements are representative:

> It (self-assessment) was a fun and great experience filled with great impact (S11).
>
> With self-assessment, I became more interested and encouraged in speaking classes (S18).
>
> It (self-assessment) made me feel like there's nothing that I can't do or learn if I just work on it and focused on it (S3).

The above quotations suggest that participants had favourable opinions and experiences with self-assessment. This positive response can also be inferred from learners' voices; namely, whenever learners spoke about their experience with self-assessment and how self-assessment affected their motivation for learning, learners' voices became enthusiastic and cheerful. Learners' motivation could be connected to learners' awareness of their role in the learning process and the significance of self-assessment in terms of improving their language progress and abilities.

The analysis of the interview data further showed that learners demonstrated awareness of how self-assessment helps them to take an active role in their learning. Learners expressed their awareness of their own active role and engagement in the learning process, several

times drawing a comparison between their role in traditional speaking assessment and their own self-assessment. For example, as one participant stated, “I liked self-assessment more (than the usual speaking assessment) because you can participate in your own progress, and it can help you notice the actual progress rather than just listen to what the teacher says about you” (S2). Several participants also explained how the self-assessment put them in charge of and allowed them to take an active role in their own educational development:

> I really liked the fact that I am the one who search for a way to improve myself after knowing my strong points and mistakes, and knowing and understanding the criteria rather than just taking a lesson after a lesson without me trying to put what I have learned in action (S1).

> I found that using the recordings and sheets encourage me to do a better job planning and preparing myself through taking notes and thinking of ways to finish the tasks in better quality (S21).

> By doing it (assessment) yourself you can know the way you prefer studying or working so you first know what the task or goal you have to achieve then you start thinking of different ways to achieve that goal and choose one to use, and afterwards you can see the results yourself and analyse them to get yourself further knowledge (S11).

In the statements above, the participants explained their experience as active learners. The participants described monitoring their performance; identifying their strength and weaknesses; and planning, preparing, and searching for ways to overcome weaknesses and correct mistakes. These statements thus suggest that self-assessment can support learners in developing a clear picture of their level, including their strengths and weaknesses, and in identifying gaps that they can overcome.

The data also suggested that, along with learners’ awareness of how self-assessment enables them to participate actively in their own learning, learners showed awareness that self-assessment requires practice. Participants observed that successful self-assessment requires practice:

> I wasn’t quite sure if this method would be effective in any way because when I tried to assess myself I can’t tell what is right and what is wrong, but after a few

> attempts and relistening to my recording and knowing and understanding the standards and the point the I need to focus on I got the hang of it better and I can judge myself and see where I was lacking and work on myself (S11).
>
> In the beginning, I was not sure if I'm doing it (self-assessment) correctly but I started to use the sheet to assess myself then re-listen to the audio and check my evaluation. At first, I notice that I was hard on myself and would do it again but starting from the fourth task I stopped doing this because after the third task I realised that I can now accurately evaluate myself and start thinking of new ways to improve (S34).

The statements above show that the learners had developed an awareness of how much effort, dedication, time, and practice it takes to be critical of their own performance, and pinpoint areas for improvement in their learning and the means by which these areas may be improved.

6.3.2.2. The benefits of self-assessment

The analysis of the interviews showed that all 10 learners agreed that self-assessment has a beneficial effect on several aspects, including their language performance and themselves as learners. The learners showed awareness of the role of self-assessment in improving their language performance and achievement:

> I think I still have a long way to be fluent like natives when speaking English but I am happy with how I improved through this journey (S11).
>
> I think my speaking skills improved a lot that even I myself can see the different before and after this semester (S2).
>
> I think it (self-assessment) really help with improving my speaking skill (S35).

The above quotations show that learners perceive self-assessment as having a significant influence on language proficiency and achievement. In turn, awareness of the beneficial effects of self-assessment on their language seems to significantly boost the learners' motivation and willingness to actively engage in study.

The data also suggested that engaging in their own learning through self-assessment

allowed the learners to develop confidence and boosted their sense of responsibility for future learning. Some examples follow:

> I don't think it only helped me with improving my English learning, but it also raised my confidence and helped me express myself more and not hesitate or be afraid of speaking what in my mind! Even if I'm afraid of making mistakes I don't let that fear controls and shut me anymore (S11)

> The more tasks I did the more confident I became and it made me feel like there's nothing that I can't do or learn if I just work on it and focused on it (S3)

> I became more interested and encouraged in speaking classes and week after week I notice that I sound more confident in the audio record and I speak more comfortably (S34)

The above statements indicate that the more learners are actively involved in their learning, the more likely learners are to develop confidence and self-motivation. This evidence further supports that the learners developed critical thinking dispositions, including self-confidence and courage. Notably, the students' voices became excited whenever they discussed their experience, which also supports the fact that self-assessment influenced the learners' motivation for learning.

The interviews analysis indicated that self-assessment creates opportunities for self-regulated learning. The learners, in detailing their experience with self-assessment, revealed the use of higher-order cognitive skills in Bloom's taxonomy (Anderson & Krathwohl, 2001) such as evaluation and analysing. The learners also revealed the use of metacognitive strategies including monitoring, planning, reflection, and brainstorming during speaking classes. The following examples highlight some of those skills and strategies:

> I started to have a goal plan more in my head or paper before I start talking just to get all my ideas in place and focus on the comment that I made in the previous one and every time I finish I re-listen to myself and I automatically evaluate myself like, did I stay on topic? Did I used adjectives? was my pronunciation clear? then I start to think of very week point and think how can I be better at it like for example one week I could not say some words so I thought I can ask my father's Canadian

> friend to say it to me and send it by whatsApp so I can repeat it and say it to him (S33)

The learners also revealed the use of metacognitive strategies including monitoring, planning, reflection, and brainstorming during speaking classes.

> It helped me plan more efficiency. What I mean is that now I first understand and set my goal and what I want to reach before even planning, then I ask myself questions and develop my plan according to the most effective way while putting in considerations what can go wrong and how to deal with the outcomes (S11)

> I now know that planning and organising your ideas and thoughts beforehand can really improve your speech and helps a lot with making it more professional and let you make less mistakes than you usually would (S21)

> I think they helped my quit much after observing myself, because now I think of multiple ways that can extend the meaning of my speech without making it boring, and I also try my best to minimise my grammar mistakes (S1).

In the above statements, learners described how they set goals and planned a strategy for approaching a speaking assessment. Learners also described the way they monitored, evaluated, analysed, and reflected on their performance. In addition, one learner described how she reflected on her performance and found a solution to a problem she encountered during the assessment. These responses suggest that higher-order cognitive skills and metacognitive strategies can be linked to self-assessment; specifically, self-assessment may have the potential to promote the development and application of these strategies and thus promote self-regulated learning.

Along with the development of cognitive and metacognitive skills, the interviews produced evidence of the impact of self-assessment on learners' metacognitive knowledge. Learners shared statements that indicated metacognitive knowledge:

> I think they made me see different sides and different view of my skills that I usually don't really consider (S1)

> Through the self-assessment I realised that I don't really focus on my own

> mistakes, but with it I can now notice where I lack and put extra effort when learning, for example I notice that I struggle with the sentences structure especially if it includes adjectives or adverbs so now when I'm learning I focus more on this and try to figure when to use it or what is the right structure for the sentences (S11).

> I felt responsibility, yes, more responsibility, more consciousness of my skills, and what do I need to develop, that helped me more, I knew exactly what I need to change in myself, sometimes, I felt sad, because there were many things I need to adjust, things I need to improve (S18).

The quotations above indicate an agreement among the learners of the impact of self-assessment on learners' awareness of their strengths and weaknesses. With the use of self-assessment, the learners identified their successes and struggles in their performance; the result was a higher level of awareness of strengths and weaknesses. This awareness can help learners to enhance their learning, as it indicates to learners where to focus and identify areas requiring further development. The data also suggested that the criteria included in the self-assessment proforma played a significant role in raising learners' awareness. The significance of the self-assessment proforma is discussed in more detail in the following section.

6.3.2.3. The significance of the self-assessment proforma

Another common theme that emerged from the analysis of the interview data is learners' awareness of the significance of knowing and understanding speaking assessment criteria and the relevance of these criteria. Informing learners of the speaking assessment criteria seemed to lead learners to perceive the criteria as vital to the learning process; learners came to understand that the assessment criteria could help them enhance their learning and identify areas that need improvement. This understanding is demonstrated in the following statements.

> I think it (criteria) made me see different sides and different view of my skills that I usually don't really consider them (S1).

> Everything in it (the criteria in assessment proforma) was important for a speech and I understood what I need to do to improve (S35)

> Honestly at first I thought what's the point to care about adjectives and adverbs in a speech but little did I know they are important. I didn't know the difference between them so I had to search about it therefore I can apply them in the next task (S34)

The quotations above indicate that providing learners with clear assessment criteria can encourage and help learners to take an active role in their learning process. Thus, clear criteria could also assist learners in defining their learning goals, identifying strategies for achieving those goals, and ultimately reaching the required level.

Evidence from the data analysis also suggests that all of the learners found the assessment criteria to be relevant and appropriate. Learners' positive reaction to the criteria is demonstrated in the following statements:

> I guess they do cover most of the points that you should have in order to be a good speaker (S1).

> I think all the criteria in the sheet were important and need to put in consideration (S11).

> I feel that is a complete unit, each part completes the other, I don't think we add anything else (S18).

In short, the learners agreed on the relevance and appropriateness of the assessment criteria. However, one participant did contradict herself. Though she first stated that the criteria were relevant, she then stated that she did not understand the reason for including time and topic management. That contradiction indicates that the criteria were not relevant to her: "In fact I'm not sure why should we evaluate our time management and topic management. It has nothing to do with our speaking performance and language" (S3). This participant's experience could indicate misunderstanding on the part of the learner on the importance of the two, or a lack of communication between the learner and the teacher.

6.3.3. Learners' perceptions and experiences of feedback in speaking classes

Feedback seems to be an issue on which the EFL learners held strong views. This theme concerns how learners described and viewed their experience with receiving feedback and

the sources of the feedback. The learners' views and attitudes are discussed and presented through the following two subordinate themes: perceptions of receiving feedback and source of feedback.

6.3.3.1. Perceptions of receiving feedback

The data from the interviews revealed that the EFL learners held conflicting views on the importance of receiving feedback. The learners also expressed different feelings about asking for or receiving feedback. Nonetheless, the majority of the learners understood the value of feedback in identifying mistakes and improving performance:

> I could use some feedback, because others may notice some points that I might miss so I can improve my language (S33).

> There are lots of times you don't notice the mistakes you make that's why asking my friends helped me a lot to be better (S11).

The statements above demonstrate the learners' awareness of the potential for feedback to have a beneficial effect on learning and development. However, two participants disagreed with the other learners, expressing negative perceptions of the importance and the role of feedback. Their negative perceptions are demonstrated in the following statements:

> I like to stay strong and don't feel weak in front of others, I don't like to be judged or criticised, this idea makes me stressed, I was in situations that forced me to do mental block and perform the task without thinking, to show everyone that I really don't care while I care (S2).

> Criticism is always a negative thing, it crushes my spirit and kills the enthusiasm (S18)

The above statements, which illustrate strong negative reactions from two learners towards feedback, could be linked to previous experience with feedback. For example, the learners' use of the word "criticism" and the phrase "crushes my spirit and kills enthusiasm" could relate to a previous and dramatically negative experience with feedback. It is possible that, in past classes, the learners received feedback that did not serve the purpose of improving learning. These negative perceptions of feedback could have a detrimental effect on learners' motivation, engagement with their own learning, and willingness to use

comments to influence future learning.

The data also suggested that, although most of the learners believe in and understand the importance of feedback, the learners held varying opinions on which sources of feedback they prefer. The question of sources of feedback is discussed in the following sub-section.

6.3.4.2. Sources of feedback

The analysis of the interviews indicates that learners held differing views on who should be the provider of feedback. Learners shared their perspectives and experiences with seeking and receiving feedback from their teacher, classmates, family members, and friends. Interestingly, only one participant expressed a preference for communicating with and receiving feedback from their teacher: "I asked the teacher to clarify her comments and feedback each time I presented, and I asked her to give me tips on how to overcome the issues" (S34). The data suggested that preferences for feedback from certain sources could be linked to previous experiences with receiving feedback from those sources. The following quotations demonstrate how past experiences could influence preferences:

> Most of the time even if someone said something wrong the teacher wouldn't pay attention to it and after it's done she doesn't give feedback, only saying good job. So you don't really know if you did a good job or not because in the end all that actually mattered were the grades from the speaking (S18).

> Normally the teacher does not point at what mistakes you made and most of the time she focuses on the PowerPoint presentation and what information I display in it instead of focusing on how to improve our performance (S21).

The above statements illustrate how learners' past experience of a teacher's involvement and the quality of feedback provided by the teacher influenced their willingness to seek and engage with feedback from their teachers. The following examples show another way in which past experiences can influence learners' preferences:

> Just the idea to ask my teacher or my classmates was a big no for me, like what if they make fun of me? Or am I just giving them the wrong first impression? What if they thought I will never improve? I had no confidence whatsoever …one time I pronounced a word wrong while talking with the teacher and she laughed at me

> then fixed it. But then from time to time when I use that word she would comment about me doing a good job that I didn't mispronounced it (S33).

> I once made a mistake in a group project with my classmates and from that day till the end of the semester they avoided working with me because I ruined it (S3).

Given the significance of saving face in Saudi society, the above statements exemplify how students' fear of embarrassment influenced their feedback choices from teachers and peers. The following example illustrates an additional way in which prior experiences might affect a learner's motivation to seek and engage with feedback:

> It is competitive in the class and everyone want to be better and have higher grades than others so they will tell you wrong things to see you fail (S35).

This statement demonstrates the influence of negative peer competition on leaners from seeking and engaging with feedback.

All the above statements illustrate how prior experiences with feedback influenced learners' willingness to seek and receive feedback from their instructor and peers. The analysis of the interview data further showed that learners preferred receiving feedback from their family members, and friends. This preference is demonstrated in the following quotations:

> I kept asking my friends who aren't in the same class as me how did I do and asked them about their thoughts on it ... I was more comfortable asking my friends than asking the teacher because I'm shy (S11).

> I didn't want to ask my classmates or my teacher, rather, after the recording, I asked my friends to hear the recording and assess my performance, what are the areas that I might have been mistaken, which might need more concentration (S33).

> That's why I prefer asking my close friends for help even though they are not taking the course with me (S3).

The statements above demonstrate learners' preferences to receive and engage with feedback from their friends outside the classroom rather than from their teacher or classmates. The following examples show students' preferences to receive and engage with

feedback from others rather than their teachers.

> To be honest I'm not that close with any of my classmates so I wouldn't feel comfortable letting them judge my work. The same goes towards my teacher, I don't want her to have the wrong impression about me. That's why whenever I wanted feedbacks I would ask my family, since I'm around them all the time they can notice my improvements more than anyone else, and I would ask them to listen to the recordings and they would tell me their honest feedback (S1).

> I always had my mother support in my learning so she can now help me with my speaking and give me notes that I can work on (S35).

These statements suggest that students would rather get feedback from their family members than from their teachers and classmates.

6.4. Conclusion

This chapter analyses the data that were gathered using three different tools. First, quantitative data were collected through audio recordings of the first and final speaking activity. Qualitative data were collected by the second and third tools, namely self-assessment proformas and semi-structured interviews. The quantitative data were analysed using descriptive statistics and a non-parametric test, namely Wilcoxon Signed Rank. The qualitative data collected by interviews were analysed with two different methods and processes of analysis. Data from the self-assessment sheets were used to trace learners' usage of self-regulated learning strategies. The data from the semi-structured interviews were analysed using thematic analysis, particularly IPA stages for thematic analysis.

The results of the quantitative data analysis revealed a statistically significant positive change in learners' speaking grades prior to and after the self-assessment intervention. The chapter presents EFL learners' perspectives and experiences of self-assessment under three themes: perceptions and experiences of traditional speaking assessment, self-assessment in speaking classes, and feedback in speaking classes.

The first theme showed that learners' perceptions of assessment are directly influenced by their experience with traditional assessment in speaking classes. The learners believed that assessment was primarily for summative purposes and had no purpose in assisting their

learning in any form. Learners identified this belief as a source of anxiety and as a reason for shifting their concentration from improving in the language to finding ways to achieve higher grades. The second theme revealed that learners' perceptions of self-assessment were affected by their experience with traditional assessment in speaking classes. The learners noted the importance of self-assessment and its function in helping them to become active participants in the learning process, along with its benefit in developing their use of higher-order cognitive skills and metacognitive skills. Additionally, learners expressed awareness that self-assessment needs to be practised. The third and last theme revealed an awareness of the importance of feedback and its positive impact on learning and development among learners. However, the perspectives of the learners on their preferred sources of feedback (teacher, family, friends, and classmates) were diverse.

The following chapter discusses the findings presented in this chapter along with the findings from Chapter 5 and situates these findings within the context of the broader literature in order to address the research questions.

Chapter 7 : Findings and Discussion

7.1. Introduction

This chapter provides a thorough analysis and discussion of the research outcomes and significant findings of this study. The outcomes and findings are derived from the descriptive and inferential analyses of quantitative data conducted through SPSS software (see Chapters 5 and 6) and from the thematic and IPA analysis of qualitative data gathered through self-assessment proformas and semi-structured interviews (see Chapter 6). The present chapter also synthesises these outcomes and findings together with those from prior studies with a view to answering the research questions. However, as discussed in Chapter 3, there are still some notable gaps in the literature, particularly with regards to the relationship between self-assessment and self-regulated learning, self-assessment and critical thinking, and self-regulated learning and critical thinking. These gaps have added another layer of difficulty to the study's theoretical foundation. Due to these gaps in the literature, I was driven to reflect on emerging findings in the context of a wider body of literature, such as disciplinary culture and feedback, in order to enrich the interpretation of this study's results.

This chapter is organised and discussed chronologically in accordance with the research questions:

1) How do EFL students perceive the implementation of self-assessment in speaking classes, especially in relation to their speaking language achievement?
2) What influence does self-assessment in speaking have on EFL students regarding:
    a. their self-regulated learning in English language speaking?
    b. their critical thinking skills?
3) What is the association between students' perceptions regarding the influence of self-assessment in speaking on their self-regulated learning and their critical thinking?

Moreover, as discussed in Chapter 4 this study adopted a mixed methods design to enable a thorough overview of the relevant and important issues from various sources of data in order to present a narrative closer to reality. Recurrent themes have emerged through cross-case analysis, and at least more than half of the participants in the interviews voiced those themes (as discussed in Chapter 6). Notwithstanding, this chapter's discussion of evidence

takes into account the prominence of the individual's voice, using individual statements when appropriate.

7.2. How do EFL students perceive the implementation of self-assessment in speaking classes, especially in relation to their speaking language achievement?

The findings of this study reveal that a number of variables shaped and influenced EFL learners' perceptions of the implementation of self-assessment in speaking classes: learners' prior experience with traditional speaking assessment, learners' motivation and willingness to self-assess, learners' awareness of assessment criteria, and learners' perceptions and experiences of feedback. The findings also reveal the impact of learners' perceptions and experiences of self-assessment in speaking classes on their language achievement. This research is original and significant in that it contextualises and discusses these ideas within the context of the EFL field in Saudi Arabia, an area where there remains very little published research in this field. Chapter 8 will then address the issues and implications emerging from this chapter.

7.2.1. EFL learners' prior experience with traditional speaking assessment

Chapter 2 argued that language instructors in higher education institutions in the Saudi context still often adhere to traditional approaches to language teaching and assessment. Evidence emerging from Chapter 6 suggests that learners' perceptions and experiences of traditional speaking assessment influence their experience and perceptions of self-assessment in speaking classes. The EFL learners' interviews, as discussed in Chapter 6, indicate that in speaking classes, lecturers and language instructors carry out assessment mainly to evaluate and measure learners' progress and performance instead of supporting their learning. For example:

> Normally the teacher does not point at what mistakes you made and most of the time she focuses of the PowerPoint presentation and what information I display in it instead of focusing on how to improve our performance (S21).

The student clearly believes that the teacher's main concern is to get through the curriculum, "she focuses of the PowerPoint presentation", rather than on how to support the student's learning and "the teacher does not point at what mistakes you made". The use

of the word “instead” suggests that this student would prefer a stronger focus on student progress. This statement indicates that the student considered that content knowledge was more significant to the teacher than language development. The following statements reveal that the students felt assessment was conducted just to evaluate and grade their language proficiency.

> When we do the speaking, the teacher ask everyone to do it as fast as possible so she can have the time to grade everyone before the class ends (S33)

The student clearly believes that the teacher’s main concern is finishing the task at hand, which is “grading everyone before the class ends” rather than focusing on each student. The use of the word “grade” suggests an emphasis on the summative purpose of an assessment of measuring learners’ progress.

> Most of the time even if someone said something wrong the teacher wouldn’t pay attention to it and after it’s done she doesn’t give feedback, only saying good job. So you don’t really know if you did a good job or not because in the end all that actually mattered were the grades from the speaking (S18).

This statement suggests that the student feels neglected by the teacher, and that the learner is not looking for words of encouragement or approval, but instead for the support of the teacher, which was not present even when making mistakes. This student also emphasised that the grades were the primary focus.

It seems that lecturers and language instructors are mainly focused on summative purposes of assessment. This approach is strongly teacher centred. It is thus in line with Al-Seghayer’s (2015) findings that language assessment in the Saudi context has merely summative purposes due to a dominating exam-oriented culture. This culture drives language teachers to stick to traditional teaching and assessment methods; under this system, teachers focus on preparing students for examinations. The result is a concentration on the topics and information covered by the tests. This adherence to traditional teaching and assessment methods could be influenced by the context and discipline.

Jessop and Maleckar (2016), Lindblom-Ylänne et al. (2006), Neumann et al. (2002), Neumann (2001) and Smeby (1996) have all argued that the disciplinary culture plays a

significant role in forming learning and teaching. This argument views teaching and assessment, as an area of practice, based on the main disciplinary categories identified in the work of Becher (1987): hard pure (denoting pure sciences like chemistry and biology), soft pure (denoting the humanities and social sciences), hard applied (denoting technology-related studies like engineering), and soft applied (denoting applied social sciences like law). For example, Lindblom-Ylänne et al. (2006) argued that disciplines and contexts influence the adoption of teaching and assessment approaches; thus, teachers with exposure to a variety of contexts may adapt their pedagogical approach according to the context. In addition, Lindblom-Ylänne et al. linked the adaption of pedagogical approach with the teachers' discipline. Similarly, Jessop and Maleckar (2016) suggested that the discipline may have an impact on the policies, practices, perceptions, and purposes of assessment. Indeed, the influence of the field and the context also emerged in Almosaa's (2021) study, which found that teachers in English language courses in a variety of Saudi universities used unified assessment practices and purposes, with minor alterations. This decision among teachers, in turn, could be attributed to the nature of English language courses in the foundation year. This issue is discussed in Chapter 8 as a proposal for future research.

The findings in Chapter 6 also indicate that the focus on exams and the summative purposes of assessment pressured learners and led them to focus on improving their exam results. For example:

> In speaking classes, you don't really care if you did a good job or not because in the end all that actually mattered were the grades from the speaking (S11).

> Most of the English classes that I have participated in and like specifically speaking classes, I always end up caring about my grades at that moment and I rarely come out of that class with benefits (S1).

These statements indicate that, as a result of the assessment's purpose, EFL learners concentrated on memorising scripts and information for their speaking assessment, studying only for the test, instead of actually improving their English language. In short, this form of assessment ultimately limited students' learning and resulted in superficial or surface approaches to learning. Biggs (1987) identified superficial or surface approaches as those emerging when learners perform assessment activities only in order to earn a grade, a mark, or a certificate; learners relying on these approaches are not driven by an interest in

or a wish to deepen comprehension of the subject matter. This finding is consistent with Heywood (2000), in argument that traditional assessment techniques, namely tests, promote surface approaches to learning. Additionally, this finding also aligns with Simonson et al. (2000) finding that test-focused learning often relies only on lower-order cognitive abilities in Bloom's taxonomy (Anderson & Krathwohl, 2001), memorising in particular, and frequently places an excessive emphasis on the final outcome, instead of higher-order ability such as analysing and evaluating.

Although the significance of assessment for summative purposes has been recognised, as discussed in Chapter 3, this emphasis on summative purposes and test-focused learning is incompatible with the primary objective of EFL instruction in KSA, which is to develop learners' intellectual, personal, and professional abilities (MoE, 2005). This adherence to exams and summative purposes of assessment could indicate that the Saudi system of education is in need for a cultural shift. Thus, greater emphasis should be placed on assessment as a means of supporting and improving students' learning as well as the learner's involvement in the process through self-assessment as an example of a practice. This suggestion among others will be addressed in Chapter 8.

Moreover, Heywood (2000) argued that traditional assessment techniques such as exams increase student anxiety. The findings from this study supported Heywood's conclusion. All learners in this study indeed considered the traditional speaking assessments, namely speaking tests, to be a source of fear and anxiety. Expressions such as "demanding", "scariest", "difficult", "stressful", and "nerve-racking" emerged when the learners talked about traditional speaking assessments. This result is in line with previous research by Mohammed (2016) showing that tests are identifies as a source of anxiety for Saudi EFL learners. It is also in line with the theory of foreign language classroom anxiety (Horwitz et al., 1986), in which test anxiety is identified as one of three components of foreign language classroom anxiety: communication apprehension, fear of negative evaluation, and test anxiety. Given that a high degree of foreign language anxiety might impede language learning (Horwitz et al., 1986), it is vital for both language instructors and students to take steps to reduce anxiety levels. This suggestion will be addressed in Chapter 8.

Additionally, learners' interviews made apparent that fear and anxiety about being assessed, as well as the value placed on grades, could have a negative impact on the level of learners' motivation to learn and make language progress. For example:

> The normal way of evaluation was never enjoyable. They were like the scariest, slowest and most stressful times. I had to memorize my answers even if I didn't understand them, because all what we ever cared about were the marks (S3).

This finding is compatible with prior research in EFL contexts by Liu and Huang (2011), Mohammed (2016) and Cayli (2020). The findings of Liu and Huang (2011), Mohammed (2016) and Cayli (2020) indicated that learners' motivations to learn are negatively correlated with the level of anxiety.

The above findings, taken together, could be seen as a positive sign of learners' awareness of their role and the role of assessment in the learning process as all learners showed dissatisfaction with this form of assessment and questioned its role in their learning process. From the interviews, it could be argued that the negative perceptions of and experiences with traditional speaking assessment could influence learners' perceptions of and experiences with self-assessment. This relationship is discussed in depth in the following subsection.

7.2.2. Learners' perceptions and experiences of self-assessment in speaking classes

Kyndt et al. (2011) suggest that traditional assessment promotes surface learning; approaches to achieving deeper learning rely, in contrast, much more on students' motivation. Self-assessment, on the other hand, is argued to be a practice that promotes deep learning and relies on students' motivation (Panadero et al., 2013; Sitzmann et al., 2010). The participants in this study displayed a high degree of enthusiasm and motivation for participation and a strong commitment to change and progress. The findings from learners' interviews in Chapter 6 show that the participants were actively engaged in self-assessment and had favourable opinions about and experiences with self-assessment in speaking classes. In the interviews, the learners described how the self-assessment empowered them to take control of actively participating in their own educational progress. For example:

> I found that using the recordings and sheets encourage me to do a better job planning and preparing myself through taking notes and thinking of ways to finish the tasks in better quality (S21).

> By doing it (assessment) yourself you can know the way you prefer studying or working so you first know what the task or goal you have to achieve then you start thinking of different ways to achieve that goal and choose one to use, and afterwards you can see the results yourself and analyse them to get yourself further knowledge (S11).

Moreover, learners' interviews suggested that the learners were eager for a shift away from the traditional approaches of learning and assessment towards new and improved approaches. For example, as one participant stated, "I liked self-assessment more (than the usual speaking assessment) because you can participate in your own progress, and it can help you notice the actual progress rather than just listen to what the teacher says about you" (S2). Additionally, whenever speaking about their experiences with self-assessment and how self-assessment affected their motivation for learning, the learners became enthusiastic and cheerful, using expressions such as "comfortable", "confident", "liked", "happy", "interested", and "encouraged". This result is in line with previous research showing that learners' motivation is essential to actively conduct self-assessment (Panadero et al., 2013; Sitzmann et al., 2010).

Additionally, Panadero et al. (2013) argued that students' motivation and willingness to utilise self-assessment tools – for example, rubrics and scripts, which are complex self-assessment scaffolding tools – relate to students' perception of the tools' use. The more valuable these self-assessment tools are to students, the more motivated students will be to utilise them (Panadero et al., 2013). Notably, some participants expressed admiration for the intervention and self-assessment proforma. For example, one participant wondered if she could use the proforma in the writing classroom to assess her writing. Thus, the learners' positive attitude and motivation could be attributed to their awareness of their active role in the learning process and their recognition of the significance of self-assessment in terms of improving their metacognitive abilities, as discussed in Section 7.3.1, and language.

Moreover, the findings in Chapter 6 also reveal that learners' knowledge of the speaking assessment criteria through the self-assessment proformas led them to the conclusion that awareness and understanding of the speaking criteria is crucial to the learning process. For example:

> I really liked the fact that I am the one who search for a way to improve myself

> after knowing my strong points and mistakes, and knowing and understanding the criteria rather than just taking a lesson after a lesson without me trying to put what I have learned in action (S1).

From learners' descriptions in the interviews, it was clear that learners understood the value of the assessment criteria which could help them identify areas for development and, ultimately, improve their learning. With clear assessment criteria, the participants in the study felt encouraged to actively engage and take ownership of their own learning. The learners interviewed in this study explained how the speaking criteria aided them in identifying their strengths and weaknesses, setting their learning objectives, developing methods for achieving those objectives, and eventually attaining the required standards. These findings support the argument of Panadero et al. (2013) and Panadero and Romero (2014), namely that the use of precise, well-defined, and well-understood criteria and standards of performance is critical to helping learners assess their own work to determine whether their performance meets standards and expectations, an action which can add to the accuracy of self-assessment. These findings are also consistent with those of Andrade and Valtcheva (2009), Panadero et al. (2013), and Panadero and Jonsson (2013), as these prior studies also found that complex self-assessment scaffolding tools, such as rubrics and scripts, could assist beginners by providing clarity on what constitutes quality performance.

Nonetheless, as Panadero, Brown, et al. (2016), Panadero and Alonso-Tapia (2013), and Cao and Nietfeld (2005) argued, providing feedback to students is necessary for them to develop into more accurate self-assessors. Without feedback that enables students to draw comparisons between their own and others' perceptions of their work, self-assessment appears to be highly dependent on individual qualities and differences (Cao & Nietfeld, 2005; Panadero & Alonso-Tapia, 2013; Panadero, Brown, et al., 2016). The following sub-section discusses learners' perceptions and experience of feedback.

7.2.3. Learners' perceptions and experiences of feedback in speaking classes

As discussed in Chapter 3, without feedback, self-assessment would be a constructivist approach which places the burden for learning on the shoulders of the learners. Such an approach leaves learners with full responsibility for their own learning. This approach leads to superficial implementation of self-assessment. Rather, educators should place self-

assessment within the sociocultural approach, in which learners develop in a social setting rather than in isolation. Within a sociocultural perspective, self-assessment should be interactive and dialogic in nature in order develop students into autonomous and self-regulated learners. Nicol and Macfarlane-Dick (2006) and Nicol (2010) argued for the conceptualisation of feedback as a dialogical two-way process involving teacher-student and/or peer-to-peer interaction as well as each learner's active involvement through self-feedback. Additionally, Panadero, Brown, et al. (2016) argued that, through teacher and peer evaluation or feedback, novice self-assessors could become aware of the possibility of inaccuracy in their assessment.

The evidence presented in Chapter 6 indicates that learners had mixed views on the value of feedback. The majority of participants, 8 out of 10 participants, were aware of the value and the power of feedback as a necessary component in the learning process, especially when it came to recognising strengths and weaknesses and improving performance. For example:

> I could use some feedback, because others may notice some points that I might miss so I can improve my language (S33).

> There are lots of times you don't notice the mistakes you make that's why asking my friends helped me a lot to be better (S11).

Indeed, the literature stresses the role of feedback as a highly effective strategy for increasing student achievement (Ahea et al., 2016; Boud & Molloy, 2013; Brooks et al., 2019; Hattie & Clarke, 2018; Hattie & Timperley, 2007; Nicol & Macfarlane-Dick, 2006; Sadler, 2010; Wisniewski et al., 2020). The data revealed that seven of the learners who participated in the study seemed to value all three aspects of feedback when providing self-feedback, especially the last two, based on the Hattie and Timperley (2007) model of feedback. Those three aspects are, first, *feed up*, which entails providing information to students and/or teachers about the learning objectives to be met. Second, *feed back* entails informing students and/or teachers about learners' accomplishments in relation to some predetermined standard or previous performance after a comparison of their current work with previous work. Third, *feed forward* entails providing information to students and/or teacher to inform future efforts and targets based on current work or performance.

Two learners, however, disagreed with the other learners, expressing unfavourable

attitudes towards the importance of feedback. The following statements reflect their unfavourable perceptions:

> I like to stay strong and don't feel weak in front of others, I don't like to be judged or criticized, this idea makes me stressed, I was in situations that forced me to do mental block and perform the task without thinking, to show everyone that I really don't care while I care (S2).

> Criticism is always a negative thing, it crushes my spirit and kills the enthusiasm (S18)

The word "criticism" and the phrase "crushes my spirit and kills enthusiasm" emerged when these two students talked about feedback. These responses were related to a previous and dramatically negative experience with feedback. The two learners' interviews suggest that their negative perceptions of feedback had a detrimental effect on their motivation, engagement with their own learning, and willingness to use comments to influence future learning, as will be discussed later in this section.

Moreover, the data revealed that the participants demonstrated a consciousness of the significance of two factors. The first was the learners' active participation in the feedback process. The second was the ways in which dialogue and discussion – whether internal, through self-monitoring, or with others, family and friends – could affect and enhance learners' understanding of feedback. As Carless (2016) observed, feedback as dialogue contributes to learners' engagement and empowerment. Feedback is an effective learning tool that allows learners to make sense of information from a variety of sources and to apply that information to their own work or learning practices in order to improve (Carless, 2016). Nicol and Macfarlane-Dick (2006), Nicol (2010), Carless et al. (2011), and Boud and Molloy (2013) argued that feedback should be conceptualised as a dialogical two-way process involving teacher-student and/or peer-to-peer interaction as well as the learners' active involvement through self-feedback. Feedback as dialogue is argued to promote learners' active involvement with feedback (Price et al., 2011), and self-regulated learning (Winstone et al., 2017). It is remarkable; however, that although learners recognised the importance of feedback and dialogue for their learning process, they declined to receive or engage in dialogic feedback with their teacher.

The findings from Chapter 6 revealed that prior experiences with feedback affected

learners' willingness to seek and engage with feedback from their teachers and peers. From the interviews, three reasons emerged related to learners' previous experience with feedback. The first reason was learners' prior experience with teacher's engagement and the quality of the feedback. This reason was by far the most common, with seven out of ten participants mentioning it. For example:

> Most of the time even if someone said something wrong the teacher wouldn't pay attention to it and after it's done she doesn't give feedback, only saying good job. So you don't really know if you did a good job or not because in the end all that actually mattered were the grades from the speaking (S18).

> Normally the teacher does not point at what mistakes you made and most of the time she focuses of the PowerPoint presentation and what information I display in it instead of focusing on how to improve our performance (S21).

The second reason was students' fear of embarrassment given the importance of saving face in the Saudi culture. This reason was identified by six learners. For example:

> Just the idea to ask my teacher or my classmates was a big no for me, like what if they make fun of me? Or am I just giving them the wrong first impression? What if they thought I will never improve? I had no confidence whatsoever …one time I pronounced a word wrong while talking with the teacher and she laughed at me then fixed it. But then from time to time when I use that word she would comment about me doing a good job that I didn't mispronounced it (S33).

As shown by the statement, the student preferred not to ask for or receive feedback from teachers or classmates owing to previous negative experiences with it and embarrassment.

The third and last reason, noted by four learners, was the negative competition between classmates, which discouraged students from seeking feedback.

> It is competitive in the class and everyone want to be better and have higher grades than others so they will tell you wrong things to see you fail (S35).

Notably, only one participant expressed a preference for communicating with and receiving feedback from their teacher: "I asked the teacher to clarify her comments and feedback each time I presented, and I asked her to give me tips on how to overcome the

issues" (S34). All the other participants, nine out of ten, preferred not to receive feedback from their teachers or their classmates, preferring instead feedback from family members and friends or peers outside the classroom. This preference is demonstrated in the following examples:

> I kept asking my friends who aren't in the same class as me how did I do and asked them about their thoughts on it ... I was more comfortable asking my friends than asking the teacher (S11).

> To be honest I'm not that close with any of my classmates so I wouldn't feel comfortable letting them judge my work. The same goes towards my teacher, I don't want her to have the wrong impression about me. That's why whenever I wanted feedbacks I would ask my family, since I'm around them all the time they can notice my improvements more than anyone else, and I would ask them to listen to the recordings and they would tell me their honest feedback and points my strengths and weaknesses (S1).

These findings are not consistent with prior research, which instead indicates EFL learners' awareness of the value and importance of teachers' feedback in their learning process (Putri & Munir, 2021). Putri and Munir (2021) findings reveal that all EFL learners considered teachers' feedback to be vital to the learning process. Additionally, based on Gamlem and Smith's (2013) classification of feedback as positive and negative feedback, Putri and Munir (2021) also found that learners recognised the value of teachers' positive feedback, i.e., feedback that expresses approval of performance, accomplishment, or effort, and specifies how the work could be improved.

Gamlem and Smith (2013) argued that feedback for approval is essential for building strong teacher-student relationships, but is not particularly beneficial for enhancing work. They also argued that feedback for approval could be either a source of motivation or a message that hinders motivation and learning. Putri and Munir (2021) findings suggest that EFL learners perceive feedback for approval with expressions such as "good", "excellent", and "OK" as positive feedback, which could increase their motivation in learning English. However, the findings from this research indicate learners perceive the expression "good job" as insufficient and useless in feedback. This perception could relate to how the expression "good job" functions in the Saudi culture, where a speaker may use the expression as a filler before changing the subject without necessarily implying that

something has been accomplished well.

Interestingly, learners' interviews also revealed that learners found feedback and comments from their family and friends outside the classroom to be more valuable than the feedback from their teacher. The learners expressed their dissatisfaction with the quality of feedback provided by their teacher. The data indicate that students believe the feedback they received from their teachers, if any, was not constructive. In contrast, learners believed that the feedback they received from their family and friends was supportive, motivating, and beneficial to their language development. For example:

> They (the family) would tell me their honest feedback and points my strengths and weaknesses (S1).

> I always had my mother support in my learning so she can now help me with my speaking and give me notes that I can work on (S35).

From a sociocultural perspective, VanPatten and Williams (2015) indicated that learners' positive or negative perspectives of learning are shaped by active participation in the cultural and linguistic contexts formed by their family and peers as well as by educational establishments. Learners' preferences in receiving feedback from family and friends outside the classroom instead of from the teacher could indicate a cultural issue as well as a weak teacher-student relationship. The statements "she laughed at me" and "I don't want her to have the wrong impression about me" suggest that a barrier is discouraging the learners from connecting and communicating with their teacher. This barrier could indicate a traditional passive relationship within the classroom between the learners and the teacher. Al-Wassia et al. (2015) argued that this relationship is the result of a cultural perception in Saudi culture that only the teacher has knowledge mastery. Wu et al. (2015) suggested that, under this type of relationship, learners are unlikely to develop their sense of responsibility, their ability to cooperate, and their critical thinking and problem-solving abilities. Thus, it is critical to consider how to overcome this barrier and build a stronger relationship between students and their teachers. Chapter 8 addresses this suggestion.

Moreover, learners believed that their family and friends would not criticise their performance, or at least would not only highlight their weaknesses, but would provide positive feedback and information that could improve their performance. Gamlem and Smith (2013) argued that the effectiveness of teacher and peer feedback is contingent upon

the involvement of ideas for improvement and honesty. This argument is perhaps one explanation for the participants' preference for feedback from family and friends outside the classroom instead of from their teacher. In short, the participants believed that the feedback from family and friends would be constructive. Additionally, the Saudi context is family-oriented, so learners tend to have close relationships with their families, which may also account for the learners' preference for feedback from family. In other words, the learners may feel more at ease seeking and engaging with feedback from family members. Consequently, I would argue that, depending on the context, family feedback may be just as valuable to the learning process. Chapter 8 will address this issue as a potential contribution of this research.

As noted in Chapter 6, the interviews also indicated that the learners viewed peer feedback as critical to improving their English language, in agreement with the findings of Alshammari (2016). However, the findings in this research indicate that peer feedback falls into two categories: peer feedback inside the classroom (*classmate-peer feedback*) and peer feedback outside the classroom (*friends-peer feedback*). As discussed above, the participants in this research preferred to seek and engage with feedback from friends outside the classroom and had negative attitudes towards feedback from classmates. This result could emphasise the significance of the social context when providing and receiving feedback. I would argue, depending on the context, friends-peer feedback could be as valuable or even more valuable than classmate-peer feedback to the learning process depending on the context of the learning.

Moreover, evidence from learners' interviews indicated that the support received from family and friends was critical to their English language learning. These results are compatible with prior research, indicating that family influence supports and improves foreign language acquisition (Gardner et al., 1997; Han, 2007).

The findings discussed in this section highlight the importance of caution when providing feedback, especially negative feedback as suggested by Ahea et al. (2016). Learners should have positive feelings about feedback as a means to motivate them to actively engage with feedback (Piccinin, 2003). Notably, Goodrich (1996) argued that teachers need to learn and develop the ability to provide learners with not only opportunities to practise self-assessment but also critical feedback about learners' self-assessment and task performance. Additionally, it is crucial to emphasise, however, that providing quality academic feedback may be difficult if students are not trained or if the criteria of assessment are not

understood (Gamlem and Smith, 2013), and the same could apply for the family members. Therefore, it is critical to explore and consider various means of overcoming this barrier. Chapter 8 will later address this suggestion.

7.3.4. The impact of learners' perceptions and experiences of self-assessment in speaking classes on their language achievement.

The findings presented in Chapter 6 revealed a positive relationship between self-assessment and EFL language achievement. The findings from the quantitative data (learners' audio recordings) presented in Chapter 6, indicate that, after the self-assessment intervention, learners' speaking grades improved. Learners' results revealed a statistically significant positive change and improvement in their speaking. These findings align with the findings from the participants' interviews. All 10 learners believed that self-assessment had a beneficial influence on their language proficiency and achievement. For example,

> I think my speaking skills improved a lot that even I myself can see the different before and after this semester (S2).

> I think it (self-assessment) really help with improving my speaking skill (S35).

The learners' awareness of the beneficial impacts of engaging in self-assessment seemed to considerably increase their confidence, motivation and willingness to actively participate in their learning.

> I don't think it only helped me with improving my English learning, but it also raised my confident and helped me express myself more and not hesitate or be afraid of speaking what in my mind! Even if I'm afraid of making mistakes I don't let that fear controls and shut me anymore (S11).

> I became more confident because I saw my progress and I become more excited and encouraged to do more tasks and improve more even when we did not have a speaking task I created one to practise more (S2).

The literature on the relationship between self-assessment and achievement offer several relevant insights. In a recent critical review of research on learners' self-assessment, Andrade (2019) reported that, without exception, the studies included in the review, which also included two meta-analyses (Graham et al., 2015; Sanchez et al., 2017), revealed a

positive relationship between self-assessment and achievement. Moreover, the present study's findings are in line with findings from research in the field of EFL learning in Saudi Arabia targeting other language skills; writing (Alshammari, 2016; Nalliveettil & Mahasneh, 2017) and reading (Rabiah, 2020); and general language improvement (Qasem, 2020).

7.3. What influence does self-assessment in speaking have on EFL students regarding their self-regulated learning and critical thinking skills.

7.3.1. The impact of learners' perceptions and experiences of self-assessment in speaking classes on their self-regulated learning.

As previously discussed in Chapter 3, self-assessment shares a significant theoretical relationship with self-regulated learning (Andrade & Cizek, 2010; Panadero & Alonso-Tapia, 2013; Panadero et al., 2018). Panadero, Jonsson, et al. (2016) indicated that the key premise in the relationship between self-assessment and self-regulated learning is that when a teacher offers opportunities for students to engage with and utilise self-assessment in the classroom, these opportunities may assist students in developing their ability for self-assessment and, therefore, their self-regulated learning skills. However, despite the strong theoretical connection between self-assessment and self-regulated learning, Yan (2020) remarked that empirical studies, included in Brown and Harris (2013) and Panadero et al. (2017) in this topic, have been far from definitive.

On the basis of the evidence in Chapters 5 and 6, self-assessment involves learners in the assessment process and increases learners' sense of control of their learning and their use of self-regulatory skills. Evidence from the pre and post SRL-SRS survey, presented in Chapter 5, indicated that self-assessment had a positive impact on learners' perceptions on the use of their self-regulatory skills and self-efficacy before and after the intervention. Chapter 5 evidence indicates that the intervention had a beneficial effect on students' perceptions of their planning skills and processes. The interviews reported in Chapter 6 support these findings, with learners explaining and discussing their experiences with self-assessment. Specifically, the interviews revealed that learners became aware of the importance of planning and setting goals at the beginning of each speaking task, and of how that step can help in improving their language and performance. For example,

> It helped me plan more efficiency. What I mean is that now I first understand and

> set my goal and what I want to reach before even planning, then I ask myself questions and develop my plan according to the most effective way while putting in considerations what can go wrong and how to deal with the outcomes (S11)

Additionally, learners' awareness of the assessment criteria and requirement assisted them in setting goals and planning for the task completion. Thus, learners began setting goals and planning as a first step of completing or performing any speaking task. In alignment with learners' perceptions regarding their planning skill, evidence collected from the learners' self-assessment proformas demonstrated that all learners used the proformas to plan for their speaking showing improvement in their planning tactics throughout the eight weeks. This result could indicate that the intervention of self-assessment and the use of the self-assessment proformas had a positive impact on the first phase of self-regulated learning (i.e. the forethought phase).

Similarly, evidence from Chapter 5 indicates that the intervention had a significant and positive influence on learners' views of their self-monitoring abilities and processes. The data revealed highly positive attitudes and self-confidence among learners regarding their ability to monitor and check their own performance; there was a statistically significant positive change post the intervention. In Chapter 6, learners explained how they were self-monitoring their speaking performance, using the self-assessment proforma according to the criteria and standards of assessment. Learners also explained how they made sure that they were on the right track while completing the task. For example,

> I really like to use the sheets because it help me focus on the requirement and criteria and it helps me to focus and try my best when I plan the task and check if I'm doing great work when I'm doing it and after I finish (S21).

In accordance with the learners' perceptions of their self-monitoring skill, evidence gathered from the learners' self-assessment proformas revealed that all learners used the proformas to monitor their speaking, indicating that their monitoring skills improved over the course of the eight-week period. This result could imply that the self-assessment intervention also had a beneficial effect on the second phase of the self-regulated learning as well (i.e. the performance phase).

Finally, evidence from the pre and post survey revealed a statistically significant positive change in learners' perceptions of their evaluation processes prior to and after intervention.

Additionally, evidence from the survey showed a statistically significant positive change in learners' perceptions of their reflection processes prior to and after the intervention. These results are supported by findings from the participants' interviews. Namely, the participants reported in their interviews higher willingness and engagement in evaluating and reflecting on their work against the assessment criteria and standards during the intervention. Learners explained how their metacognitive knowledge and awareness of their strengths and weaknesses during the evaluation process helped them find solutions to overcome their weaknesses during reflection, and eventually led them to enhance their learning and performance. For example,

> Through the self-assessment I realized that I don't really focus on my own mistakes, but with it I can now notice where I lack and put extra effort when learning, for example I notice that I struggle with the sentences structure especially if it includes adjectives or adverbs so now when I'm learning I focus more on this and try to figure when to use it or what is the right structure for the sentences (S11).

This result could indicate that the self-assessment intervention had a beneficial effect on the third phase of the self-regulated learning (i.e. the evaluation or appraisal phase).

Taken together, these results indicate that the self-assessment proforma activates the usage of self-regulatory skills and engages the learners in all three phases of self-regulated learning: the forethought phase, in which tasks are analysed, plans are made, and goals are established; the performance phase, which involves the usage of learning and monitoring strategies; and the evaluation or appraisal phase, in which students reflect on and evaluate their learning achievements. These findings suggest that the intervention was beneficial in boosting students' use of self-regulated learning skills. The findings thus link back to the literature, supporting the theoretical relationship between self-assessment and self-regulated learning; that is, that engaging students in self-assessment can develop their self-regulated learning skills (Andrade & Cizek, 2010; Andrade, 2019; Panadero & Alonso-Tapia, 2013; Panadero et al., 2018; Panadero, Jonsson, et al., 2016). Additionally, these findings are consistent with the findings reported in Andrade's (2019) critical review and Panadero et al.'s (2017) meta-analytic review, namely that self-assessment interventions are beneficial for students' self-regulated learning skills. Nonetheless, the effect size in Panadero et al. (2017) research varied from small to medium, according to the tool used for self-regulated learning, which is not consistent with the effect size in this research. Moreover, in the literature on self-assessment and self-regulated learning within the field

of EFL learning, these findings align with those of Kahrizi et al. (2014) from language schools that self-assessment has a statistically significant positive impact on self-regulated learning. However, the study by Kahrizi et al. (2014) study only reports statistical significance; as discussed in Chapter 5, statistical significance is proportional to the sample size, so the occurrence of statistical significance alone does not guarantee the presence of an effect. The findings of this research provide empirical evidence of the relationship between self-assessment and self-regulated learning within the field of EFL learning and particularly during speaking classroom at HEI. Thus, I would argue that the self-assessment proforma could be considered as a complex self-assessment scaffolding tool to support student learning. Chapter 8 will address this as a potential contribution of this research.

Moreover, as discussed earlier in this chapter, self-assessment relies on students' motivation (Panadero et al., 2013; Sitzmann et al., 2010). Students' motivation, in turn, refers to the extent to which a student makes an effort and maintains a focus on studying in order to attain success (Saeed & Zyngier, 2012). Within research on motivation in learning, learners' self-efficacy is among the most important aspects of motivation (Linnenbrink & Pintrich, 2002). In addition, learners' motivation is positively affected by their self-efficacy (Zimmerman, 1995). In other words, learners' effort is associated with learners' perceptions of their ability to perform a task and accomplish predefined objectives. Thus, self-efficacy is likely to have a significant impact on the notion of self-regulated learning and self-assessment, as previously discussed in Chapter 3.

Evidence presented in Chapter 5 from the pre and post SRL-SRS survey indicate a statistically significant positive change in learners' effort and self-efficacy prior to and after the intervention. The findings revealed a statistically significant positive change in learners' perception of the level of dedicated effort put into their work before and after the intervention. The findings also revealed a statistically significant positive change in learners' perception of their self-efficacy before and after the intervention. Learners' interviews, as shown in Chapter 6, revealed that the learners' active engagement and involvement with self-assessment in their speaking classes encouraged them to take on a more active role in their own learning; taking on that more active role, in turn, enhanced the level of effort the learners put into their work. As the learners saw the results, namely improvement, the learners came to believe in their ability to overcome difficulties. For example, "I became more confident because I saw my progress and I become more excited and encouraged to do more tasks and improve more even when we did not have a speaking

task I created one to practise more" (S2). The findings of the impact of the self-assessment intervention on the EFL learners' self-efficacy in HEI are similar to those of Baleghizadeh and Masoun (2013) in language school. Namely, Baleghizadeh and Masoun (2013) found that EFL learners' self-efficacy level increased significantly as a result of consistent use of self-assessment.

7.3.2. The impact of learners' perceptions and experiences of self-assessment in speaking classes on their critical thinking skills.

Learners who acquire and develop critical thinking abilities often continue to use such abilities well into later in life and transfer those abilities into academic and professional success (Murawski, 2014). By developing critical thinking, learners tend to broaden their worldview and improve their ability to negotiate critical choices in school and in life (Murawski, 2014; Phan, 2010). One of the arguments in Chapter 3 was that providing learners with self-assessment opportunities could enhance their critical thinking. Evidence from the Watson-Glaser II Critical Thinking Appraisal prior to and after the self-assessment intervention presented in Chapter 5 shows a statistically significant difference between learners' critical thinking total grades prior to and after the intervention. The data indicate a statistically significantly positive improvement in learners' critical thinking, which aligns with the findings of (Kahrizi et al., 2014). Kahrizi et al.'s study found that learners' critical thinking increased significantly after a self-assessment intervention in comparison to their results before the intervention. Nevertheless, Kahrizi et al. (2014) research reports only statistical significance; as noted in Chapter 5, statistical significance is related to sample size, so the existence of statistical significance does not ensure the presence of an effect.

The data presented in Chapter 5 also indicated development in each of the five critical thinking skills after the intervention. First, the evidence reveals a statistically significant positive improvement in learners' ability to make inferences about the degrees of truth or untruth of a statement. Second, the evidence reveals a statistically significant positive improvement in learners' ability to recognise underlying assumptions or presuppositions. Third, the evidence reveals a statistically significant positive improvement in learners' deductive ability, as measured by asking them to determine whether or not specific facts were definitely true. Fourth, the evidence reveals a statistically significant positive improvement in learners' interpretive ability, as measured by having the learners evaluate evidence and determine if generalisations or conclusions were appropriate. Lastly, the

evidence reveals a statistically significant positive improvement in learners' ability to differentiate between strong and relevant arguments and weak or irrelevant arguments.

Moreover, along with the development of learners' critical thinking skills, evidence from learners' interviews revealed that self-assessment enabled the learners to gain or improve two of the critical thinking dispositions, namely confidence and courage, and increased their sense of responsibility to support future learning by participating in their own learning.

7.4. What is the association between students' perceptions regarding the influence of self-assessment in speaking on their self-regulation and their critical thinking?

Dickinson (1987) and Zimmerman (2002) argued that self-regulated learning and critical thinking are interrelated ,because a self-regulated person who is critical in their thinking is able to learn more efficiently and readily in a shorter amount of time than someone who lacks these qualities. Ku and Ho (2010b) argued that metacognitive strategies can elicit actions in pupils that help them to monitor and manage their mental processes, and that the employment of metacognitive methods can often be as a significant element influencing critical thinking.

Evidence from Chapter 5 indicates that, prior to the intervention, the relationship between critical thinking and self-regulated learning skills was not statically significant. The relationship between learners' critical thinking and the overall perception of self-regulated learning was weakly negative. The relationship between critical thinking in relation to the self-regulated learning skills of planning, self-monitoring, evaluating, putting forth effort, and self-efficacy was weak; the one exception was reflection, with which critical thinking had a weak positive relationship. The correlation between critical thinking and self-regulated learning skills changed after the intervention, yet the relationships remained statistically insignificant except for that with planning skills. Learners' critical thinking had a medium-strength positive relationship with planning, reflection, and self-efficacy and a weak positive relationship with self-monitoring, evaluation, and effort. These findings are not consistent with those of Kahrizi et al. (2014). Although both studies indicate statistically significantly positive improvement in learners' critical thinking and self-regulated learning, the correlation in this study is not statically significant as opposed to Kahrizi et al. (2014)

Nonetheless, even though the relationship was not statistically significant, the findings in the present study indicate that self-regulated learning and critical thinking do share a medium-strength positive relationship. This relationship suggests that an improvement in learners self-regulated learning ability could lead to improvement, to some extent, in their critical thinking skills. In other words, the more learners apply and improve their self-regulatory skills, the more critical thinking skills are likely to develop. This finding supports the arguments of Dickinson (1987), Zimmerman (2002), and Ku and Ho (2010b). Given that both critical thinking and self-regulated learning are high-order skills, this relationship seems more feasible. Thus, the findings in this study seem to indicate that the contributing role of the relationship between self-regulated learning and critical thinking might also be valid in the Saudi context. Chapter 8 addresses this finding as a potential contribution of this research and as an area for further research.

7.4. Conclusion

This chapter presented a comprehensive analysis of the research findings in light of the existing literature in order to address the research questions. Several themes emerged from the findings in response to the three research questions. The first question sought to identify EFL learners' perceptions of self-assessment in speaking classes. The findings indicate that self-assessment in speaking classes was seen favourably by language learners. The learners in this study expressed admiration for the intervention and the assessment tool, which influenced their motivation and willingness to learn and self-assess. Learners also expressed admiration for how self-assessment allowed them to be aware of and understand the assessment criteria, and helped them engage with and improve their learning by being able to identify their strengths and weaknesses and areas for language development.

The findings discussed in this chapter suggest that learners' positive perceptions and experiences of self-assessment are mainly influenced by their negative prior experience with traditional speaking assessment. The findings in this study indicate learners' eagerness to change from traditional speaking assessments that dominate language courses in higher education institutions in Saudi Arabia. The learners expressed that traditional speaking assessments were a source of anxiety and how they are driven to focus on grades instead of improving their language, thereby driving students to memorise scripts for their speaking.

This chapter also discussed learners' perceptions and experiences of feedback, as it plays a vital role in self-assessment and learning. The learners that participated in this study recognised the significance of feedback in the learning process, as well as the value of dialogue and discussion in feedback and its role in understanding and using feedback to improve their learning. However, the learners described preferring to engage in feedback with their family and friends outside of the classroom rather than their teachers and classmates based on previous experiences with feedback, which is considered to be a key finding in this study.

The second research question aimed to investigate the impact of self-assessment in the speaking classroom as perceived by learners. This impact was firstly discussed in relation to learners' self-regulated learning skills, namely planning, self-monitoring, evaluation, and reflection, along with their effort and self-efficacy. The impact was also discussed in relation to learners' critical thinking skills, namely their ability to infer, recognise assumptions, deduce, interpret, and evaluate arguments. Finally, the third research question sought to investigate the relationship between learners' perceptions of their self-regulated learning and their critical thinking. This question aimed to test (Ku & Ho, 2010b) theory that a relationship exists between these two factors. The findings in this chapter reflect how critical thinking abilities are more likely to develop with the development of self-regulatory skills.

This study's novelty and significance lie in its contextualisation and discussion of these key concepts in the EFL field in Saudi Arabia. Furthermore, a number of key findings contribute to the study's novelty. These findings concern learners' preference for receiving feedback from family and friends, the classification of peer feedback into classmate-peer feedback versus friends-peer feedback, and the self-assessment proforma as a complex self-assessment scaffolding tool. These ideas are discussed further in Chapter 8.

Chapter 8 : Conclusion

8.1. Introduction

This research has investigated the impact of self-assessment on self-regulated learning, critical thinking, and language achievement in an EFL context in Saudi Arabia from learners' perspectives. The first and second research questions aimed to explore, in terms of EFL learners' perceptions, the influence of self-assessment on EFL learners' self-regulated learning, language achievement, and critical thinking. The third research question concerned the relationship between EFL learners' perceptions of self-regulated learning and critical thinking.

The previous chapters, in seeking to address these research questions, have provided a thorough analysis and discussion of the findings in light of the existing literature. This chapter now provides a reflection on this thesis and its contributions to the study's field and context. The chapter then provides some recommendations derived from the findings of the study. This chapter concludes by noting the limitations of the study and recommending areas for further research.

Prior to addressing the aforementioned points, it is vital to acknowledge that this research is merely a snapshot of the process of development in the higher education (HE) sector in Saudi Arabia, where policies, practices, and theories are in a constant state of transformation in the process of reaching Saudi Vision 2030. Additionally, this research only investigated a small part of the participants' learning experiences. The results of this research are not intended to be generalised to a larger population of undergraduate students.

The results of this research are not intended for generalisation to a larger undergraduate students' population. These results are intended instead to stimulate questions, opportunities, and possibilities for teachers, researchers, policy makers, and curriculum designers who are interested in the research findings and aim to investigate the findings' applicability to their own context.

8.2. Critical reflection on the research aim and its contribution

This study investigated the perspectives of Saudi EFL students learning English in a preparatory foundation year programme at a Saudi Arabian higher education institution throughout one academic semester. This research is significant in light of the Saudi Arabian government's substantial investment and commitment in conceiving, securing, and integrating curriculum transformations in Saudi Arabia's educational institutions in the process of achieving Saudi Vision 2030. Given that assessment is a fundamental component of the curriculum, research such as the present study is vital to empowering learners and supporting English language learning. A critical reflection on the findings of this research regarding the aim and research's contributions is presented in the following sub-sections.

8.2.1. Learners' support for a shift from assessment of learning to assessment for learning

The Saudi Ministry of Education aims to provide an extraordinary educational system that uses "a modern curriculum focused on rigorous standards in literacy, numeracy, skills, and character development" (Saudi-Vision-2030, 2016, p. 38). Accordingly, assessment in Saudi Arabia should shift to a social-realist perspective of knowledge, a foundation that embraces how the social aspect of assessment can enhance learning. However, according to Case (2015), the institutions, curricula, and modes of teaching and assessment will need to undergo considerable structural and cultural transformation in order to create an undergraduate experience that will allow for a shift in students' agency and in cultural and social structure. This shift will include abandoning the 'survival of the fittest' approach in favour of assuming that the majority of students accepted for the programme should succeed, specifically via intensive engagement with themselves, others, and the discipline knowledge. It requires a transition from assessment as judgement to assessment for learning.

This transition, however, is not free from challenges. Achieving reform and transformation in any culture is likely to be gradual, with success subject on the adaptability of the context (Creanza et al., 2017). Allport et al. (1954) suggested that EFL learners develop their attitudes through direct interactions with various sociocultural elements, including history, culture, and L1 context. Research in the context of Saudi Arabia indicates the challenge of obtaining support for the adoption of learner-centred strategies (Alyami, 2016; Fatany,

2009). However, the findings in this research indicate that the learners, who are used to traditional assessment and teacher-centred learning, are now more receptive to new strategies. Comments from the students suggest that the students have positive attitudes towards opportunities that would give them the chance to actively participate in and be involved in their own education. This change in attitudes may be influenced by the context in Saudi Arabian culture which is beginning to welcome changes in many aspects in life. Formerly a more traditional and conservative society, Saudi Arabia is today striving to become part of a larger global community, and is also working to develop its knowledge economy as part of its 2030 vision. Despite the country still being in the early phases change and development, this shift in students' perspectives demonstrates a change in how education is considered. This might be considered as evidence that both the Saudi Ministry of Education and the students desire to reform the assessment system. Therefore, Saudi Arabia's implementation of self-assessment in the EFL curriculum in higher education might be a step towards reaching the objective of reforming the education system, of which skills and character development form a key part.

8.2.2. Using self-assessment to promote deep learning, self-regulated learning, and critical thinking

Kyndt et al. (2011) argued that learners' use of learning approaches is affected by assessment demands. Kyndt et al. (2011) suggested that traditional assessment promotes the use of approaches that lead to surface learning. The findings in this study seem to indicate that, due to the nature and purpose of assessment in speaking classrooms, and due to how significant grades were to the learners, the EFL learners adopted the surface learning approach of memorisation for their speaking assessment. Self-assessment, in contrast, led learners to begin to adopt deeper learning approaches, i.e. understanding and reflection, for their speaking. This evidence supports E. Panadero et al. (2013) argument that self-assessment is a practice that promotes deep learning. The shift between learning approaches might also be considered to be reflective of how students adapt to suit the requirements of assessment. Biggs (1987) identified superficial or surface approaches as those approaches that emerge when learners perform assessment activities solely to earn a grade, a mark, or a certificate; learners who rely on these approaches are rarely motivated by an interest in or a desire to gain a deeper understanding of the subject matter. Additionally, as discussed previously in this research, self-assessment is an effective practical instructional strategy to improve learning and the quality of learning. Self-assessment, if carefully enacted, was found to be effective in the work of Panadero et al.

(2017) and Brown and Harris (2014). However, this relationship needs to be explored in different educational contexts, as the findings in research exploring the relationship are often presented as universally valid. The findings in the Saudi EFL context in higher education, a context where there is no previously published work, indicate that self-assessment offers opportunities for language learners to use and develop their self-regulated learning skills. Therefore, the adoption of self-assessment in the EFL curriculum in Saudi Arabia could enable students to engage in deep and meaningful learning and thus improve the quality of learning.

Moreover, as argued by Dickinson (1987) and Zimmerman (2002), self-regulated learning and critical thinking are interrelated; thus, an increase in learners' ability to self-regulate their learning might result in an improvement in their critical thinking skills. The results of this research also suggest that the use of self-regulated skills may be a key signal of learners' improved critical thinking skills, and vice versa. This association between self-regulated learning and critical thinking may play a significant role in the Saudi context in relation to the Saudi government's 2030 vision and the goal of developing a knowledge economy, which emerges with increasing usage of knowledge-based outputs obtained via intellectual capabilities; such knowledge-intensive tasks demand independent thinkers who are capable of creating and solving problems. This finding may also indicate the need to employ and explore different strategies and create opportunities for the development of learners' self-regulated learning and/or critical thinking skills, which might lead to improvement in both areas to empower the learners for their future. This also requires further research regarding the level of the impact and improvement of different strategies on learners' self-regulated learning and critical thinking skills.

This research proposed and tested the self-assessment proforma as a complex self-assessment scaffolding tool, as a means of improving the quality of learning through engaging students in deep and meaningful learning and the development of the learners self-regulated learning skills and, thus, their critical thinking skills. The self-assessment proforma may also be presented as an artefact that mediates learning, which is discussed in the following section.

8.2.3. Self-assessment intervention and proforma as mediating artefacts.

Looking through the lens of activity theory by Leont'ev (1978), the present research also proposes the self-assessment proforma as an artefact in the idea of mediation by artefacts.

The fundamental idea, according to this theory, in the activity theory is that artefacts, such as tools, methods, and processes, have a mediating function in activities. As described by Patchen and Smithenry (2014), mediating artefacts provide transformational learning opportunities in the classrooms. In this research, the self-assessment proforma mediates learning and the development of self-regulated learning skills. Through the use of the self-assessment proforma as an artefact is considered as a means by which the learners can concentrate their thinking on specific aspects of their learning, which is, in this study, the development of their language and the use and development of their self-regulatory skills. The findings indicate that the proforma clarifies assessment expectations and standards, helping learners to understand the assessment expectations and standards as well as to determine whether their performance met the standards and expectations. Additionally, the self-assessment proforma activated the use of self-regulatory skills and engaged the learners in all three phases of self-regulated learning: the forethought phase, in which tasks are analysed, plans are made, and goals are established; the performance phase, which involves the use of learning and monitoring strategies; and the evaluation or appraisal phase, in which students reflect on and evaluate their learning achievements. Therefore, the self-assessment proforma could indeed act as a complex self-assessment scaffolding tool, and the findings indicate that it was beneficial in boosting students' use of self-regulated learning skills, fostering deep learning, and improving the quality of learning.

Moreover, building on the idea of mediation by artefacts, the work of Conole (2008) and Conole (2012) on learning design has used the term *mediating artefacts* (MS) to refer to the different forms of representation of any learning activity. According to Conole (2008), learning activities "can be 'codified' into a number of different forms of representation, which each foreground different aspects of the learning activity" (p. 4). Conole (2008) and Conole (2012), furthermore, argues that MS have the potential to assist in making informed decisions in terms of implementing activities. Through identifying mediating artefacts in a learning activity or an intervention, practitioners are better able to more accurately comprehend how learning activities are represented and how these artefacts may be utilised to facilitate new design. Case study or narratives, iconic presentations or diagrammatic, vocabularies, and models are some of the examples of different forms of representation or mediating artefacts for learning activities as proposed by Conole (2008) and Conole (2012). Accordingly, the present research is proposed as *case study mediating artefacts,* describing the specifics of a pedagogical intervention (Conole, 2008), namely the use of a self-assessment proforma during speaking classes to promote self-regulated learning. Additionally, the diagram in Figure 8-1 is proposed as a *diagrammatic mediating*

artefacts for the intervention as a means of providing a concise summary of the essential aspects of the intervention and the use of the self-assessment proforma. Diagrammatic or iconic presentations, according to Conole (2008), highlight the key aspects of the intervention and their connections and an indication of structure or process.

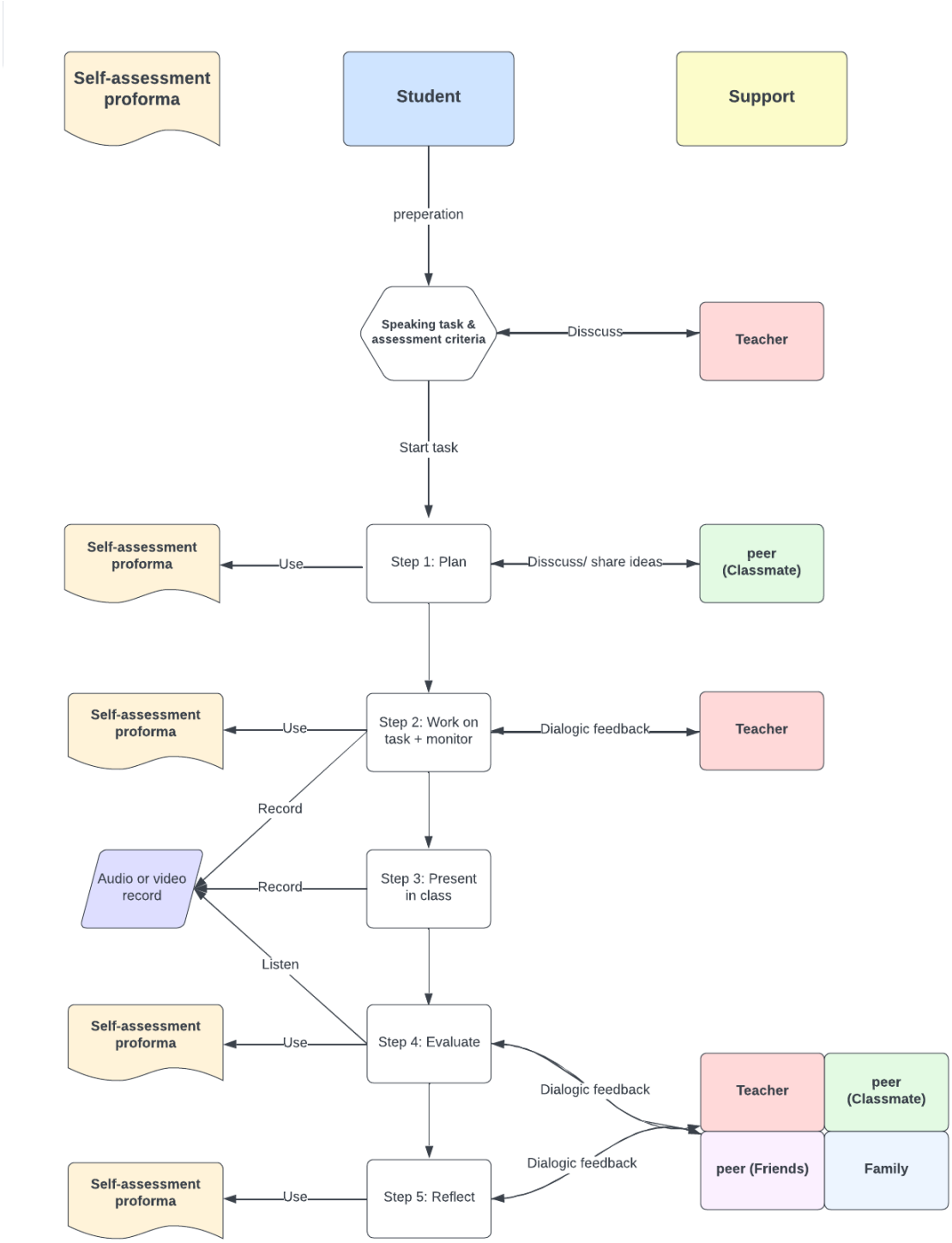


Figure 8-1 diagrammatic mediating artefacts for the self-assessment intervention

Moreover, Conole (2008) argued that the mediating artefacts might be joined with guidelines, key points or Frequently Asked Questions (FAQs). Therefore, the following aspects need to be taken into account for effective use or repurpose of the mediating artefacts:

- **Transparency** involves sharing precise, well-defined, and well-understood criteria in the self-assessment proforma. Clear standards of performance are critical to helping learners assess their own work and determine whether their performance meets expectations (Panadero et al., 2013; Panadero & Romero, 2014). This necessitates engaging learners in discussing these goals and criteria to ensure a common understanding and, when possible, to prompt learners to choose their learning goals, which might increase learners' motivation (Hayward, 2013).

- **Feedback** involves engaging learners in dialogic feedback with their teachers and peers. Dialogic feedback, according to Carless (2016), enhances learners' active engagement and understanding of feedback.

- **Modelling** or **exampling** involve providing an example of a self-assessment proforma for the students to review. Though not vital, according to Zimmerman and Kitsantas (2005), models of instruments can increase learning potentials and benefits.

Moreover, as discussed previously in this research, feedback plays a significant role in achieving the anticipated learning benefits in self-assessment. Therefore, it is vital to acknowledge that many of the issues discussed in this study in terms of feedback are consistent with international findings on feedback. However, the next section discusses a distinctive feature in relation to feedback that has emerged in the Saudi culture.

8.2.4. The significance of feedback from family and friends

As discussed in Chapter 3, self-assessment, from a sociocultural viewpoint, should be dialogical and participatory in order to promote autonomy and self-regulated learning (Nicol, 2010). Feedback, as argued by Nicol and Macfarlane- Dick (2006) and Nicol (2010), should be a two-way, dialogical process including teacher-student and/or peer-peer interaction in addition to the learner's active participation via self-feedback. Through teacher and peer feedback, novice self-assessors can become aware of the possibility of

inaccuracy in their assessment (Panadero, Brown, et al., 2016). As the study's evidence indicates, learners were aware of the importance and efficacy of feedback as an integral part of the educational process. The learners also valued dialogue feedback as it could influence and improve their understanding of feedback. Yet, despite learners' acknowledgement of the value of feedback and discussion to their learning process, they did not seek or engage in dialogic feedback with their teacher. Interestingly, the participants in this study preferred and chose to receive and engage with feedback from family and friends outside the classroom rather than from the teacher. More interestingly, the participants in this study did not seek or engage with their family or friends for emotional support or feedback, as in the findings of Massri (2017) in the Saudi context in which students stressed the role of the family relationship and highlighted the importance of family support in English language learning, in terms of emotional support in particular. The students in the current study preferred to engage with family rather than with their teacher and to discuss their speaking performance, including their strengths and weaknesses and ways for improvement. This preference poses questions regarding the teacher-student relationship and the parent-student relationship.

To understand students' feedback-related preferences and choices in this study, I drew on research on students' engagement with feedback and parent involvement in higher education to inform an understanding of the aspects that influence students' feedback decisions. Within the literature on student engagement, Handley et al. (2011) and Carless (2019) have identified the stage prior to engagement with feedback (i.e. the pre-engagement stage) as vital to learners' engagement with feedback. Handley et al. (2011) identified this stage as "readiness-to-engage", namely, students' willingness to engage with feedback. According to Handley et al. (2011), learners' willingness to engage or disengage with feedback is determined by a number of variables, including their past experiences with it. Likewise, learners' prior experiences with feedback had great importance to learners' engagement with feedback in a study by Carless (2019) using the 3P (presage, process, product) model. This model was based on the Biggs 3P models of learning and teaching (Biggs, 1993, 1999 as cited in Carless (2019). The 3P model addresses three aspects of the feedback experience for learners: "presage or prior experiences; processes of engaging with feedback; and products, namely the likely outcomes and impact of the processes" (Carless, 2019, p. 53). According to Carless (2019), the impact of the feedback varied according to the influence of the presage and process stages. I have focused here on the presage stage of this model to understand students' choices and preference of feedback in this study. This stage includes the following several factors: prior experience with

feedback, the competences required to effectively engage with feedback, and the motivation to utilise the feedback to develop. This stage also includes the teaching context, which involves course design, teaching inputs, learning activities, and assessment design, and it includes relational factors, including the course atmosphere and the relationships between participants (Carless, 2019, p. 53).

As Handley et al. (2011) and Carless (2019) found, learners' prior experience with feedback has a significant impact on their engagement with feedback. In addition, according to VanPatten and Williams (2015), learners' positive or negative attitudes towards learning are shaped by learners' active engagement in the cultural and linguistic settings produced by their family and peers as well as by educational institutions. The findings from the current study indicate that learners' refusal to engage with feedback from their teachers and classmates tends to be rooted in negative prior experiences with feedback and the traditional passive relationship between the teacher and students. This finding emphasises the influence of learners' prior experiences on their engagement with feedback and ultimately their learning process. Nonetheless, several important questions remain regarding the participants' reasons and preferences for engaging with feedback from family. These questions concern, in particular, the characteristic of parent involvement during higher education and the Saudi context. However, despite the significance of parental involvement in higher education and the interaction between students and their parents, little study has taken place in the context of Saudi Arabia.

The body of literature on parent involvement in the higher education context has generally acknowledged the role of the parents as a core component in student development (Lowe & Dotterer, 2018; Sax & Wartman, 2010). Yet Sax and Wartman (2010) highlighted that despite the academic community's acknowledgement of parental involvement as a legitimate strategy for promoting student success, there is limited consensus on its theoretical and practical definitions and characteristics. In a recent study, Lowe and Dotterer (2018) define parental involvement as a "multi-dimensional construct composed of three distinct involvement strategies: parental support giving, parent-student contact, and parental academic engagement"(p. 36).

Looking through the lens of the theoretical underpinnings of parental involvement, and through Bowlby's (1988) developmental theory of attachment in particular, evidence from this study suggests that relationships between students and parents enable students' confidence to proceed in their learning and offer support in the form of dialogic feedback

when necessary. This might indicate that the learners are aware of the significance of learning in a social context; and that, due to their prior experience with feedback from teachers and classmates and their relationship with their family, learners described preferring the support of their family over that of their teacher. Investigating the characteristics and impact of parental involvement is, therefore, necessary to understand students' feedback choices and identify the best practices that parents and family may employ to support learners. Moreover, the findings of this study indicate that peer feedback could be divided into two distinct categories: peer feedback inside the classroom (classmate-peer feedback) and peer feedback outside the classroom (friends-peer feedback). Further research should investigate the characteristics of each type of feedback and identify the best practices for each type to be most effective. This distinction, however, may be a result of learners' prior negative experience with peer-feedback. Al-Khairy (2013) contends that peer pressure might contribute to the demotivation of language students. Thus, learners' attitudes towards classmate-peer feedback may improve if learners study in a friendly classroom environment in which the learning setting fosters trust and support.

Taken together, the findings from this research indicate that learners' feedback decisions are subject to their emotions. This is highlighted in the work of Immordino-Yang (2015), that emotions become a characteristic of the cognitive skills such as learning, motivation, memory, and decision making, thereby affecting these skills. Therefore, understanding students' emotions and the roots and purposes of these emotions regarding feedback and their teacher-student relationship and the parent-student relationship could be used to find solutions.

Stepping back from the findings in this research on feedback, what is really interesting is that it would be expected that students see their English language teachers as experts in terms of language teaching and learning, as those who have the content knowledge in terms of language teaching and learning, and the pedagogical knowledge in terms of approaches to learning and an understanding of the curriculum and what is required for progression in examinations. Therefore, it would be expected that students in terms of their own success, would want to access this knowledge and have this professional feedback. However, the findings in this research suggest that students prefer to have that pedagogical support to actively support their learning from their family and friends rather than their teacher. This issue poses the following important questions:

Is pedagogical support, including feedback from an "expert", devalued in the Saudi culture?

Do the cultural associations of family and friends indicate that, if the students have feedback and advice from their family and their teacher, are the students more likely to take advice from their family or friends rather than the teacher?

Is this decision only influenced by the students' emotions, or does it have a wider prevalence in the Saudi culture?

Thus, further research is required to answer these questions.

Meanwhile, the findings from this research indicate that higher education institutes should take into consideration the significance of family and friends-peer feedback on EFL learning. Nonetheless, it is vital to highlight the likelihood of inaccuracy and ineffectiveness in feedback from family and friends. The educational background and language proficiency of family and friends might influence their ability to provide accurate and effective feedback. Using technology, as well as facilitating teacher and peer feedback, particularly for the new generations in Saudi Arabia, might offer an answer for this dilemma (Al-Hariri & Al-Hattami, 2017). Additionally, the recent interest in "technology-enhanced feedback" indicates considerable potential for technology-enhanced feedback to eliminate some of the hurdles that may hinder face-to-face dialogue between learners and lecturers, hence increasing the use of feedback (Pitt & Winstone, 2020). Henceforth, the creation of an online platform for dialogue feedback between family, friends, and learners under the supervision of the teacher, as well as between classmates, teachers, and other learners, might enhance the likelihood that all learners will participate in meaningful and effective dialogue to support their learning. Further research is needed to investigate the effectiveness of such technology-enhanced feedback.

8.3. Recommendations from the study

Based on the findings of this study, this section presents the following recommendations for educational policy making, practice, and future research.

8.3.1. Recommendations for educational policy and higher education institutes

The evidence collected from this research regarding the use and impact of self-assessment to improve the quality of learning and develop learners' LLL may motivate education decision makers and higher education institutes to implement self-assessment at higher education institutes, particularly in the EFL context. The following is, therefore, recommended to decision makers.

Recommendation 1: Reduce reliance on assessment of learning and increase the use of assessment to support learning.

> One of the findings in this study was that currently, assessment is mainly used for grading the learners, and self-assessment in comparison offers the benefits of improving learning, empowering learners, and developing learners' self-regulated learning and critical thinking skills. It is, thus, recommended to encourage higher education institutions to adopt complex forms of self-assessment such as the self-assessment proforma, which is proposed in this study, rubrics, or scripts to develop self-regulated learners and effective lifelong skills and strategies, instead of attempts to adopt weak forms of self-assessment, such as check lists. Such a shift could also improve the state of language learning in the educational system.

Recommendation 2: Conducting workshops and training sessions on self-assessment for the learners.

> Another finding in this research was that the Saudi EFL learners seem eager to shift away from traditional assessment, and they show interest and willingness to self-assess. It is thus recommended to encourage higher education institutions to build a culture of support and ongoing development for learners through conducting workshops and training sessions on self-assessment. The aim is to build and develop a clear and profound understanding of self-assessment's rationale, benefits, practices, and processes. Learners should have spaces to meet regularly and share their ideas on practices of self-assessment, based on their actual experiences.

8.3.2. Recommendations for practice

The recommendations to improve the quality of learning and develop learners' LLL should be combined with recommendations for practice. Combining these recommendations would optimise the effectiveness of self-assessment and eliminate any potential negative results. The following is, therefore, recommended.

Recommendation 1: Create a friendly classroom environment to reduce students' anxieties and build strong relationships between students, teachers, and classmates.

Recommendation 2: Offer one-on-one counselling sessions to support students with self-assessment and provide feedback.

> One of the findings in this study was that the students preferred not to ask for, or receive, feedback from teachers or classmates owing to previous negative experiences, which included embarrassment. The following is therefore recommended: Encourage teachers to create a friendly classroom environment and build a strong relationship with their students based on trust and support. The aim is to reduce students' anxieties and fear of mistakes, facilitate active learning, and increase learners' motivation to learn. Additionally, encourage the teachers to foster better and stronger relationships among classmates. Secondly, encourage the teachers to offer one-on-one counselling sessions to provide assistance with self-assessment and appropriate support, such that the teachers can monitor and give dialogic feedback to guarantee effective implementation.

Recommendation 3: Encourage teachers to engage with the learners in dialogic feedback.

> Another finding in this research was that learners recognised the importance of feedback and dialogue for their learning process. Thus, it is recommended to encourage teachers and students to engage in more dialogue in the classroom. Teachers should also be encouraged to collaborate with learners to establish a shared comprehension for criteria of assessment, since the criteria are essential to the learning experience of the students.

Recommendation 4: Encourage the use and adaptation of the self-assessment proforma as a reliable scaffolding assessment method and a mediating artefact.

Another finding and contribution of this research was that the self-assessment proforma could serve as a complex self-assessment scaffolding tool and a mediating artefact assessment that can foster deep learning, clarify expectations, make criteria explicit, and facilitate feedback for a proper implementation of self-assessment. The following is thus also recommended:

> Recommendation 5: Ensure the implementation of self-assessment is underpinned by defined curriculum objectives.
>
> Recommendation 6: Encourage the use of audio or video recordings as models to foster an understanding of the criteria and learning expectations.

8.4. Limitations of the study and recommendations for future research

This study's limitations included the methodology and the limitations imposed by the nature and division of the English language course. This study included a self-assessment intervention and measured learners' critical thinking before and after the intervention to examine the impact of self-assessment on their learning. According to Pallant (2016), such research is enhanced by the inclusion of a control group that is not exposed to the intervention but is identical to the experimental group in all other aspects; several other factors, including the passage of time, may contribute to changes. Indeed, a control group would have helped to rule out the effect of other factors that may have influenced the results. However, due to the time constraint, the nature and division of the English language course, conducting the study with both experimental and control groups would not have been appropriate. The experimental group would have gained an advantage on assessment at the end of the academic semester; there would have been insufficient time to implement the intervention with the control group after completing the experiment with the experimental group. This limitation could be rectified with similar research in other higher education institutions in Saudi Arabia to corroborate the findings and contribute to the insights gained here. Thus, further research is recommended to overcome this limitation.

Furthermore, this study aimed to explore learners' perceptions of the implementation of self-assessment in a speaking classroom and its impact on their self-regulatory skills, critical thinking, and achievement. This study concentrated on speaking classroom as it aimed to target the gap in the literature regarding the use of self-assessment in speaking classes in KSA. However, focusing only on speaking may have created the misleading idea

that the components of academic literacy, including writing, listening, reading, and speaking, are distinct; in fact, they are not. Thus, further research is recommended to include all four language skills.

Moreover, this study used pre and post surveys and interviews as tools for data collection. Based on the findings from this study, if I were to recreate the research, one of the participants groups that I would add to this research would be language instructors as key participants to investigate their perspectives of self-assessment and feedback and explore their observations regarding the impact of self-assessment on EFL learners.

Additionally, I would include classroom observations to provide more in-depth information. Classroom observations would make it possible to investigate teacher-student relationships, peer relationships, and feedback procedures in more depth.

Future research may also investigate policymakers' perception of self-assessment. Due to the top-down structure of the Saudi educational system, all decisions and policies are accountable to the crown policymakers Lastly, future research may also involve an investigation of the English language curriculum to determine how assessment is handled currently. Such an investigation could shed light on ways in which curriculum-related concerns may be alleviated in order to promote improved student learning.

8.5. Reflection

I began my study with the strong belief that with the new generation of students, the spoon-feeding and conventional teacher-centred practices that are prevalent in Saudi English language classrooms will no longer be successful. The learners' interviews and perceptions of the traditional assessment practices support this belief. If we continue to educate and assess in the same traditional manner, we will continue seeing the same unsatisfactory language learning results. Unless a change occurs, it is unlikely that anything will improve. To engage and motivate students in the learning process, there is a need to adopt modern approaches. Consequently, encouraging and adopting an independent learning style is crucial not only for enhancing the quality of education in Saudi Arabia but also for preparing the next generation with lifelong skills that extend beyond the classroom. However, there is a strong need for evidence situated in the Saudi context before implementing any change. This is significant, as a different social and cultural context activates different mind-sets which may lead individuals to approach and

perceive independent learning differently.

Additionally, as mentioned earlier in this thesis, all forms of assessment in Saudi Arabia are mostly seen as a means of grading students. I began this research with an understanding of assessment as being either summative assessment, which serves the purpose of evaluating students' learning, or formative assessment, which supports students' learning. Nonetheless, as the research progressed, I have discovered that there is so much to assessment and that it is complex, as there are many interpretations and ideas underpinning it. Now, at the end of this research, I have recognised that the results collected by any assessment may be utilised for a variety of purposes and that the ultimate purpose of assessment is learning.

Moreover, the concept of self-regulated learning was one of the foundational pillars of this research. To me, initially, self-regulated learning was simply the idea of someone taking charge of their own learning. Nonetheless, as the research progressed, my understanding of the concept has changed and developed. This study has shown me that the development of self-regulated learning depends on the formation of skills and habits. Students, teachers, classmates, friends, and family members can all play a role in helping students to develop positive learning habits and self-regulated learning and lifelong skills. Thus, we, the community, need to support the new generation with more opportunities to learn and evolve.

BIBLIIOGRAPHY

Ahea, M., Ahea, M., Kabir, R., & Rahman, I. (2016). The Value and Effectiveness of Feedback in Improving Students' Learning and Professionalizing Teaching in Higher Education. *Journal of Education and Practice, 7*(16), 38-41.

Al Asmari, A. R. (2013). Saudi university undergraduates' language learning attitudes: A preparatory year perspective. *International Journal of Asian Social Science, 3*(11), 2288-2306.

Al-Abdullatif, A. M. (2020). Alternative assessment approaches and quality product design within web-based learning environments. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT), 15*(3), 60-74.

Al-Awaid, S. A. A. (2018). Teaching strategies in EFL environment in the secondary schools in the KSA: Evaluation and remedies. *Advances in Language and Literary Studies, 9*(2), 50-58.

Al-Hajailan, T. (2003). Teaching English in Saudi Arabia. *Riyadh: Aldar Alsawlatiah*.

Al-Haqwi, A. I., Al-Wahbi, A. M., Abdulghani, H. M., & Van der Molen, H. T. (2012). Barriers to feedback in undergraduate medical education. *Saudi Med J, 33*(5), 557-561. (Saudi Med J)

Al-Hariri, M. T., & Al-Hattami, A. A. (2017). Impact of students' use of technology on their learning achievements in physiology courses at the University of Dammam. *Journal of Taibah University Medical Sciences, 12*(1), 82-85.

Al-Issa, A. (2009). Education reform in Saudi Arabia between the absence of political vision and apprehension of religious culture and the inability of educational administration. *Lebanon, Beirut: Dar Al-Saqi*.

Al-Khairy, M. H. (2013). English as a foreign language learning demotivational factors as perceived by Saudi undergraduates. *European Scientific Journal, 9*(32).

Al-Motairi, A. D. M. (2005). *Towards improving the teaching of English as a foreign language in the Saudi context: a case study of five Saudi boys secondary schools* University of East Anglia].

Al-Musallam, A. (2007). Higher education accreditation and quality assurance in the Kingdom of Saudi Arabia. First National conference for Quality in Higher Education,

Al-Sadan, I. (2000). Educational assessment in Saudi Arabian schools. *Assessment in Education: principles, policy & practice, 7*(1), 143-155.

Al-Saloom, H. (1987). The History of Educational Movement in Saudi Arabia. *Saudi Arabia*. (Saudi Arabia)

Al-Seghayer. (2005). Teaching English in SA: Slowly but steadily changing. In G. Braine (Ed.), *Teaching English to the world: History, curriculum, and practice* (pp. 125-134). Routledge.

Al-Seghayer, K. (2011). *English teaching in Saudi Arabia: Status, issues, and challenges*. Hala.

Al-Seghayer, K. (2015). Salient Key Features of Actual English Instructional Practices in Saudi Arabia. *English Language Teaching, 8*(6), 89-99.

Al-Seghayer, K. (2022, August, 2022). Continued Concerns with Language Assessment Practices in Saudi Arabian English Education. *Education, Language and Sociology Research, 3*, 55-65.

Al-Subahi, A. A. H. (1988). A communicative-functional English curriculum for intermediate school in Saudi Arabia: a model for English syllabus design and implementation.

Al-Wassia, R., Hamed, O., Al-Wassia, H., Alafari, R., & Jamjoom, R. (2015). Cultural challenges to implementation of formative assessment in Saudi Arabia: An exploratory study. *Medical Teacher, 37*(sup1), S9-S19.

Alahmadi, N., Alrahaili, M., & Alshraideh, D. (2019). The impact of the formative assessment in speaking test on Saudi students' performance. *Arab World English Journal (AWEJ) Volume, 10*.

Alhamami, M. (2019). Learners' beliefs about language-learning abilities in face-to-face & online settings. *International Journal of Educational Technology in Higher Education, 16*(1), 1-23.

Alharbi, A., & Madhesh, A. (2018). Inclusive education and policy in Saudi Arabia. *International Journal of Education Research and Reviews, 6*(1), 946-956.

Alhojailan, M. I. (2012). Thematic analysis: A critical review of its process and evaluation. *West east journal of social sciences, 1*(1), 39-47. (West east journal of social sciences)

Allport, G. W., Clark, K., & Pettigrew, T. (1954). The nature of prejudice.

Almossa, S. Y. (2021). Assessment practices in English language courses in Saudi public universities. Collated Papers for the ALTE 7th International Conference, Madrid,

Almulhim, A. M. (2001). *An English language needs assessment of Saudi college-of-technology students with respect to a number of business sectors in Saudi Arabia*. The University of Mississippi.

Almutairi, M. M., & Hasanat, M. H. A. (2018). Impact of SRL Strategies and Academic Achievements on Recruitment of Information System Graduates in Saudi Arabia. *International Journal of Computer (IJC), 31*(1), 53-64. (International Journal of Computer (IJC))

Alotaibi, K. A. (2019). Teachers' Perceptions on Factors Influence Adoption of Formative Assessment. *Journal of Education and learning, 8*(1), 74-86.

Alshammari, M. (2016). The role of peer-and self-assessment in developing Saudi EFL learners' English writing skills. *International Journal of Education, 8*(3), 85.

Alshraideh, M. K. (2015). Pre-Service Teachers Reasons and Beliefs about Teaching as a Profession. *EDUCARE, 7*(2). (EDUCARE)

Alyami, A. (2016). *Cross-cultural studies among Saudi students in the United Kingdom* Brunel University London].

Alzaidi, H. (2011). *An investigation into the implementation of CLT in the Saudi female intermediate and secondary school context* Master‘s thesis), Newcastle University, Newcastle, UK].

Amit, A., & Sagiv, L. (2013). The role of epistemic motivation in individuals’ response to decision complexity. *Organizational Behavior and Human Decision Processes, 121*(1), 104-117.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Andrade, H., & Cizek, G. J. (2010). Students as the defi nitive source of formative assessment: Academic self-assessment and the self-regulation of learning. In *Handbook of formative assessment* (pp. 102-117). Routledge.

Andrade, H., & Valtcheva, A. (2009). Promoting learning and achievement through self-assessment. *Theory into practice, 48*(1), 12-19. (Theory into practice)

Andrade, H. L. (2018). Feedback in the context of self-assessment. In *The Cambridge handbook of instructional feedback.* (pp. 376-408). Cambridge University Press. https://doi.org/10.1017/9781316832134.019

Andrade, H. L. (2019). A critical review of research on student self-assessment. Frontiers in Education,

Andrade, H. L., & Brown, G. T. (2016). Student self-assessment in the classroom. In *Handbook of human and social conditions in assessment* (pp. 319-334). Routledge.

Andrade, M. S., & Bunker, E. L. (2009). A model for self-regulated distance language learning. *Distance education, 30*(1), 47-61.

ARG. (2002). *Assessment for learning: 10 principles*. Assessment Reform Group London.

Ariza, E. N., & Zainuddin, H. (2002). *Fundamentals of teaching English to speakers of other languages in K-12 mainstream classrooms*. Kendall Hunt.

Bachman, L. F. (1990). *Fundamental considerations in language testing*. Oxford university press.

Bachman, L. F. (2000). Modern language testing at the turn of the century: Assuring that what we count counts. *Language testing, 17*(1), 1-42.

Bachman, L. F., & Palmer, A. S. (1989). The construct validation of self-ratings of communicative language ability. *Language testing, 6*(1), 14-29. (Language testing)

Bagheri, F., & Ghanizadeh, A. (2016). Critical thinking and gender differences in academic self-regulation in higher education. *Journal of Applied Linguistics and*

Language Research, 3(3), 133-145. (Journal of Applied Linguistics and Language Research)

Bailey, C. A. (2007). *A guide to qualitative field research*. Sage Publications.

Bailey, K. M. (2003). Speaking. *Practical English language teaching*, 47-66.

Bailin, S., & Battersby, M. (2016). *Reason in the balance: An inquiry approach to critical thinking*. Hackett Publishing.

Bailin, S., Case, R., Coombs, J. R., & Daniels, L. B. (1999). Common misconceptions of critical thinking. *Journal of curriculum studies, 31*(3), 269-283.

Baki, R. (2004). Gender-segregated education in Saudi Arabia: Its impact on social norms and the Saudi labor market. *Education policy analysis archives, 12*(28), n28.

Baleghizadeh, S., & Masoun, A. (2013). The effect of self-assessment on EFL learners' self-efficacy. *TESL Canada Journal*, 42-42.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Bandura, A. (1988). Self-regulation of motivation and action through goal systems. In *Cognitive perspectives on emotion and motivation* (pp. 37-61). Springer.

Bandura, A. (1991). Social cognitive theory of self-regulation. *Organizational Behavior and Human Decision Processes, 50*(2), 248-287.

Barnes, C. J. (2006). Preparing preservice teachers to teach in a culturally responsive way. *Negro educational review, 57*.

Becher, T. (1987). The disciplinary shaping of the profession. *The academic profession*, 271-303.

Behar-Horenstein, L. S., Beck, D. E., & Su, Y. (2018). An initial validation study of the self-rating scale of self-directed learning for pharmacy education. *American Journal of Pharmaceutical Education, 82*(3).

Beishuizen, J., & Steffens, K. (2011). A conceptual framework for research on self-regulated learning. In *Self-regulated learning in technology enhanced learning environments* (pp. 1-19). Brill.

Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: principles, policy & practice, 18*(1), 5-25.

Berg, B. L. (2004). Methods for the social sciences. *Qualitative Research Methods for the Social Sciences. Boston: Pearson Education, 191*. (Qualitative Research Methods for the Social Sciences. Boston: Pearson Education)

Berry, R. (2008). *Assessment for learning* (Vol. 1). Hong Kong University Press.

Berry, R. S. Y. (2013). The assessment as learning (AaL) framework for teaching and learning: The AaL wheel.

Biggs, J. (1998). Assessment and classroom learning: a role for summative assessment? *Assessment in Education: principles, policy & practice, 5*(1), 103-110.

Biggs, J. B. (1987). *Study Process Questionnaire Manual. Student Approaches to Learning and Studying*. ERIC.

Birjandi, P., & Bagherkazemi, M. (2010). The Relationship between Iranian EFL Teachers' Critical Thinking Ability and Their Professional Success. *English Language Teaching, 3*(2), 135-145. (English language teaching)

Bixler, G. M., Brown, A., Way, D., Ledford, C., & Mahan, J. D. (2015). Collaborative concept mapping and critical thinking in fourth-year medical students. *Clinical pediatrics, 54*(9), 833-839. (Clinical pediatrics)

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education: principles, policy & practice, 5*(1), 7-74.

Black, P., & Wiliam, D. (2003). 'In praise of educational research': Formative assessment. *British educational research journal, 29*(5), 623-637.

Black, P., Wilson, M., & Yao, S.-Y. (2011). Road maps for learning: A guide to the navigation of learning progressions. *Measurement: Interdisciplinary Research & Perspective, 9*(2-3), 71-123.

Blanche, P., & Merino, B. J. (1989). Self-assessment of foreign-language skills: Implications for teachers and researchers. *Language learning, 39*(3), 313-338.

Blatchford, P. (1997). Students' self assessment of academic attainment: Accuracy and stability from 7 to 16 years and influence of domain and social comparison group. *Educational Psychology, 17*(3), 345-359.

Bloom, B. S. H., J. T; Madaus, G. F. (1971). Handbook on formative and summative evaluation of student learning.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and instruction, 7*(2), 161-186.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied psychology, 54*(2), 199-231.

Boekaerts, M., & Niemivirta, M. (2000). Self-regulated learning: Finding a balance between learning goals and ego-protective goals. In *Handbook of self-regulation* (pp. 417-450). Elsevier.

Boekaerts, M., & Pekrun, R. (2015). Emotions and emotion regulation in academic settings. In *Handbook of educational psychology* (pp. 90-104). Routledge.

Bogdan, R., & Biklen, S. K. (1997). *Qualitative research for education*. Allyn & Bacon Boston, MA, USA:.

Bol, L., & Garner, J. K. (2011). Challenges in supporting self-regulation in distance education environments. *Journal of Computing in Higher Education, 23*(2), 104-123.

Borg, S., & Edmett, A. (2019). Developing a self-assessment tool for English language teachers. *Language teaching research, 23*(5), 655-679.

Borman, W. C., Buck, D. E., Hanson, M. A., Motowidlo, S. J., Stark, S., & Drasgow, F. (2001). An examination of the comparative reliability, validity, and accuracy of performance ratings made using computerized adaptive rating scales. *Journal of applied psychology, 86*(5), 965.

Boud, D., & Falchikov, N. (2007). *Rethinking Assessment in Higher Education: Learning for the Longer Term*. Taylor & Francis. https://books.google.co.uk/books?id=bsF8AgAAQBAJ

Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & evaluation in higher education, 38*(6), 698-712.

Boyatzis, R. E. (1998). *Transforming qualitative information: Thematic analysis and code development*. sage.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative research in psychology, 3*(2), 77-101. (Qualitative research in psychology)

Braun, V., & Clarke, V. (2021). To saturate or not to saturate? Questioning data saturation as a useful concept for thematic analysis and sample-size rationales. *Qualitative research in sport, exercise and health, 13*(2), 201-216.

Brink, M. K. (2017). *Teachers' perceived understanding of formative assessment and how this understanding impacts their own classroom instruction*. Southern Illinois University at Carbondale.

Broadfoot, P., & Black, P. (2004). Redefining assessment? The first ten years of assessment in education. *Assessment in Education: principles, policy & practice, 11*(1), 7-26.

Brooks, C., Carroll, A., Gillies, R. M., & Hattie, J. (2019). A matrix of feedback for learning. *Australian Journal of Teacher Education (Online), 44*(4), 14-32.

Brown, G., & Harris, L. (2013). Student self-assessment. In J. H. McMillan (Ed.), *SAGE Handbook of Research on Classroom Assessment* (pp. 367–393). Thousand Oaks: Sage.

Brown, G. T., Andrade, H. L., & Chen, F. (2015). Accuracy in student self-assessment: directions and cautions for research. *Assessment in Education: principles, policy & practice, 22*(4), 444-457.

Brown, G. T., & Harris, L. R. (2014). The Future of Self-Assessment in Classroom Practice: Reframing Self-Assessment as a Core Competency. *Frontline Learning Research, 2*(1), 22-30.

Brown, H. (2000). *Teaching by Principles: An Interactive Approach to Language Pedagogy by H. Douglas*. New York: Longman.

Brown, H. D., & Abeywickrama, P. (2004). Language assessment. *Principles and Classroom Practices. White Plains, NY: Pearson Education*.

Brown, H. D., & Abeywickrama, P. (2010). Language assessment: Principles and classroom practices (Vol. 10). *White Plains, NY: Pearson Education*. (White Plains, NY: Pearson Education)

Brown, S., & Glasner, A. (1999). *EBOOK: Assessment Matters In Higher Education*. McGraw-Hill Education (UK).

Bryman, A. (2014). June 1989 and beyond: Julia Brannen's contribution to mixed methods research. *International Journal of Social Research Methodology, 17*(2), 121-131.

Bryman, A. (2016). *Social research methods*. Oxford university press.

Burns, A. (2009). *Doing action research in English language teaching: A guide for practitioners*. Routledge.

Butler, D. L. (2002). Individualizing instruction in self-regulated learning. *Theory into practice, 41*(2), 81-92.

Cal, A., & Tehmarn, A. (2016). Phenomenological epistemology approaches and implications for HRD research and practice. [Manchester, England: University Forum of Human Resource Development]. Manchester, England: University Forum of Human Resource Development,

Calfee, R., & Sperling, M. (2010). *On Mixed Methods: Approaches to Language and Literacy Research (An NCRLL Volume). Language & Literacy Series (NCRLL Collection)*. ERIC.

Campbell, D., DeWall, L., Roth, T., & Stevens, S. (1998). Improving Student Depth of Understanding through the Use of Alternative Assessment.

Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied linguistics, 1*(1), 1-47.

Cao, L., & Nietfeld, J. L. (2005). Judgment of Learning, Monitoring Accuracy, and Student Performance in the Classroom Context. *Current Issues in Education, 8*(4), n4.

Carless, D. (2005). Prospects for the implementation of assessment for learning. *Assessment in Education: principles, policy & practice, 12*(1), 39-54.

Carless, D. (2007). Learning-oriented assessment: conceptual bases and practical implications. *Innovations in education and teaching international, 44*(1), 57-66. (Innovations in education and teaching international)

Carless, D. (2012). Trust and its role in facilitating dialogic feedback. In *Feedback in higher and professional education* (pp. 90-103). Routledge.

Carless, D. (2014). Exploring learning-oriented assessment processes. *Higher Education, 69*(6), 963-976.

Carless, D. (2016). Feedback as dialogue. *Encyclopedia of educational philosophy and theory*, 1-6.

Carless, D. (2019). Learners' feedback literacy and the longer term: Developing capacity for impact. In *The impact of feedback in higher education* (pp. 51-65). Springer.

Carless, D., Joughin, G., & Liu, N.-F. (2006). *How assessment supports learning: Learning-oriented assessment in action* (Vol. 1). Hong Kong University Press.

Carless, D., Salter, D., Yang, M., & Lam, J. (2011). Developing sustainable feedback practices. *Studies in Higher Education, 36*(4), 395-407.

Carter, R., & McCarthy, M. (2017). Spoken grammar: Where are we and where are we going? *Applied linguistics, 38*(1), 1-20.

Carver, C. S., & Scheier, M. F. (1981/2012). *Attention and self-regulation: A control-theory approach to human behavior*. Springer Science & Business Media.

Carver, C. S., & Scheier, M. F. (2017). Self-regulatory functions supporting motivated action. In *Advances in motivation science* (Vol. 4, pp. 1-37). Elsevier.

Case, J. M. (2015). A social realist perspective on student learning in higher education: the morphogenesis of agency. *Higher Education Research & Development, 34*(5), 841-852.

Castagnaro, P. J. (2006). Audiolingual method and behaviorism: From misunderstanding to myth. *Applied linguistics, 27*(3), 519-526. (Applied linguistics)

Cayli, M. B. (2020). *An analysis of Efl learners' foreign language anxiety and their motivation to learn English* The University of Mississippi].

Chamot, A. U., & O'malley, J. M. (1994). *The CALLA handbook: Implementing the cognitive academic language learning approach*. Addison-Wesley Publishing Company Reading, MA.

Chen, Y.-M. (2008). Learning to self-assess oral performance in English: A longitudinal case study. *Language teaching research, 12*(2), 235-262. (Language teaching research)

Cheng, S.-F., Kuo, C.-L., Lin, K.-C., & Lee-Hsieh, J. (2010). Development and preliminary testing of a self-rating instrument to measure self-directed learning ability of nursing students. *International journal of nursing studies, 47*(9), 1152-1158.

Cleary, T. J., & Zimmerman, B. J. (2012). A cyclical self-regulatory account of student engagement: Theoretical foundations and applications. In *Handbook of research on student engagement* (pp. 237-257). Springer.

Coe, R. (2017). The nature of educational research. *Research methods & methodologies in education*, 5-14. (Research methods & methodologies in education)

Cohen, L., Manion, L., & Morrison, K. (2007). Research methods in education (Sixth). *Oxon: Routledge*.

Cohen, L., Manion, L., & Morrison, K. (2018). *Research Methods in Education (Eigth Edition ed.)*. New York: Routledge.

Coleman, B. J., Mason, P., & Steagall, J. W. (2012). Does a business curriculum develop or filter critical thinking? *American Journal of Business Education (AJBE), 5*(4), 409-416. (American Journal of Business Education (AJBE))

Conole, G. (2008). The role of mediating artefacts in learning design. In *Handbook of research on learning design and learning objects: issues, applications, and technologies* (pp. 188-208). IGI global.

Conole, G. (2012). *Designing for learning in an open world* (Vol. 4). Springer Science & Business Media.

Consolo, D. A. (2006). Classroom oral interaction in foreign language lessons and implications for teacher development. *Linguagem & Ensino, 9*(2), 33-55.

Coombe, C., & Canning, C. (2002). Using self-assessment in the classroom: Rationale and suggested techniques. *Karen's Linguistics Issues*.

Coombe, C. A. (1992). *The relationship between self-assessment ratings of functional literacy skills and basic English skills test results in adult refugee ESL learners*. The Ohio State University.

Coombs, A., DeLuca, C., LaPointe-McEwan, D., & Chalas, A. (2018). Changing approaches to classroom assessment: An empirical study across teacher career stages. *Teaching and Teacher Education, 71*, 134-144.

Correia, R. C. (2016). Assessing speaking proficiency: A challenge for the Portuguese EFL teacher. *e-TEALS, 7*(1), 87-107.

Creanza, N., Kolodny, O., & Feldman, M. W. (2017). Cultural evolutionary theory: How culture evolves and why it matters. *Proceedings of the National Academy of Sciences, 114*(30), 7782-7789.

Creswell, J. (2009). Research Design: Qualitative, Quantitative, and Mixed Methods Approaches 3rd Edn London: Sage.

Creswell, J. W. (2010). Mapping the developing landscape of mixed methods research. *SAGE handbook of mixed methods in social & behavioral research, 2*, 45-68.

Creswell, J. W. (2011). Controversies in mixed methods research. *The Sage handbook of qualitative research, 4*(1), 269-284. (The Sage handbook of qualitative research)

Creswell, J. W. (2012). *Research design: qualitative, quantitative, and mixed methods approaches (4th ed.). Thousand Oaks, CA: Sage Publications.*
. Thousand Oaks, CA: Sage Publications. .

Creswell, J. W., & Creswell, J. D. (2017). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage publications.

Croker, R. A. (2009). An introduction to qualitative research. In *Qualitative research in applied linguistics* (pp. 3-24). Springer. (Qualitative research in applied linguistics)

Crotty, M. J. (1998). The foundations of social research: Meaning and perspective in the research process. *The foundations of social research*, 1-256. (The foundations of social research)

Cullen, R. (1998). Teacher talk and the classroom context. *ELT journal, 52*(3), 179-187.

Daley, C., & Nisa, S. (2016). Evaluating a Student Centred Approach To Teaching And Learning On A Post-Graduate Professional Module. *Journal of Advances in Higher Education, 8*(1), 7-15.

Dann, R. (2002/2012). *Promoting assessment as learning: Improving the learning process*. Routledge.

Dann, R. (2012). *Promoting assessment as learning: Improving the learning process*. Routledge.

Darandari, E., & Murphy, A. (2013). Assessment of student learning. In *Higher Education in Saudi Arabia* (pp. 61-71). Springer.

Davis, K. A. (1995). Qualitative theory and methods in applied linguistics research. *Tesol Quarterly, 29*(3), 427-453. (Tesol Quarterly)

Denzin, N. (1978). *Sociological methods: A sourcebook*. McGraw-Hill.

Derwing, T. M., Rossiter, M. J., Munro, M. J., & Thomson, R. I. (2004). Second language fluency: Judgments on different tasks. *Language learning, 54*(4), 655-679.

Dewey, J. (1910/1997). *How we think*. Courier Corporation.

Dharma, I. P. S., & Adiwijaya, P. A. (2018). The Effect Of Problem Based Learning And Self-Assessment On Students' Writing Competency And Self-Regulated Learningm. SHS Web of Conferences,

Diab, N. M. (2010a). Effects of peer-versus self-editing on students' revision of language errors in revised drafts. *System, 38*(1), 85-95.

Diab, N. M. (2010b). Peer-editing versus self-editing in the ESL classroom.

Dickinson, L. (1987). *Self-instruction in language learning*. Cambridge, England: Cambridge University Press.

Dikel, M. (2005). *A guide to going online for self-assessment tools.* .

Dinsmore, D. L., Alexander, P. A., & Loughlin, S. M. (2008). Focusing the conceptual lens on metacognition, self-regulation, and self-regulated learning. *Educational psychology review, 20*(4), 391-409.

Do Quyen, N. T., & Khairani, A. Z. (2017). Reviewing the challenges of implementing formative assessment in Asia: The need for a professional development program. *Journal of Social Science Studies, 4*(1), 160-177.

Dunn, K. E., Rakes, G. C., & Rakes, T. A. (2014). Influence of academic self-regulation, critical thinking, and age on online graduate students' academic help-seeking. *Distance education, 35*(1), 75-89.

Duque Micán, A., & Cuesta Medina, L. (2017). Boosting vocabulary learning through self-assessment in an English language teaching context. *Assessment & evaluation in higher education, 42*(3), 398-414.

Earl, L. M. (2003/2013). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.

Earl, L. M. (2013). *Assessment as learning: Using classroom assessment to maximize student learning*. Corwin Press.

Earl, L. M., & Katz, M. S. (2006). *Rethinking classroom assessment with purpose in mind: Assessment for learning, assessment as learning, assessment of learning*. Manitoba Education, Citizenship & Youth.

Eatough, V., & Smith, J. A. (2017). Interpretative phenomenological analysis. *The Sage handbook of qualitative research in psychology*, 193-209.

Ebrahimi, M. R., & Moafian, F. (2012). Does Emotional Intelligence or Self-Efficacy Have Something to Do with High School English Teachers' Critical Thinking, Considering Demographic Information? *International Journal of Linguistics, 4*(4), 224. (International Journal of Linguistics)

Elgadal, H. A. (2017). *The effect of self-assessment on inexperienced EFL students' writing during revision* University of Birmingham].

ELI-Curriculum-manual. (2018/2019). *ELI_102_Curriculum manual*.

Ellis, R., & Ellis, R. R. (1994). *The study of second language acquisition*. Oxford University.

Elwood, J. (2006). Formative assessment: Possibilities, boundaries and limitations. *Assessment in Education: principles, policy & practice, 13*(2), 215-232.

Elyas, T. (2008). The attitude and the impact of the American English as a global language within the Saudi education system. *Novitas-ROYAL (Research on Youth and Language), 2*(1). (Novitas-ROYAL (Research on Youth and Language))

Ennis, R. H. (1969). Logic in teaching. *Englewood Cliffs, NJ: Prentice-Hall, Inc.*

Ennis, R. H. (1987). A taxonomy of critical thinking dispositions and abilities.

Ennis, R. H. (2016). Definition: A Three-Dimensional Analysis with Bearing on Key Concepts.

Epstein, R. M., Siegel, D. J., & Silberman, J. (2008). Self-monitoring in clinical practice: a challenge for medical educators. *Journal of Continuing Education in the Health Professions, 28*(1), 5-13.

ETEC. (2018). *Education and Training Evaluation Commission*.

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of educational research, 83*(1), 70-120.

Facione, P. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report).

Fahim, M., Bagherkazemi, M., & Alemi, M. (2010). The Relationship between Test Takers' Critical Thinking Ability and their Performance on the Reading Section of TOEFL. *Journal of Language Teaching & Research, 1*(6). (Journal of Language Teaching & Research)

Fan, J., & Yan, X. (2020). Assessing speaking proficiency: a narrative review of speaking assessment research within the argument-based validation framework. *Frontiers in Psychology, 11*, 330.

Fatany, S. (2009). *Saudi challenges and reforms*. Arab Center for Publications, Media & Researchs.

Fathi, J., Yousefi, L., & Sedighravesh, M. (2017). The impact of self-assessment and peer-assessment in writing on the self-regulated learning of Iranian EFL students. *Journal of Sociological Research, 8*(2), 1-16.

Filep, B. (2009). Interview and translation strategies: coping with multilingual settings and data. *Social Geography, 4*(1), 59-70. (Social Geography)

Firestone, W. A. (1993). Alternative arguments for generalizing from data as applied to qualitative research. *Educational researcher, 22*(4), 16-23.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist, 34*(10), 906.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education* (Vol. 7). McGraw-hill New York.

Fulcher, G. (2015). Assessing second language speaking. *Language Teaching, 48*(2), 198-216.

Gaffas, Z. M. (2019). Students' perceptions of the impact of EGP and ESP courses on their English language development: Voices from Saudi Arabia. *Journal of English for Academic Purposes, 42*, 100797.

Gamlem, S. M., & Smith, K. (2013). Student perceptions of classroom feedback. *Assessment in Education: principles, policy & practice, 20*(2), 150-169.

Gardner, J. (2012). *Assessment and learning*. Sage.

Gardner, R. C., Tremblay, P. F., & Masgoret, A. M. (1997). Towards a full model of second language learning: An empirical investigation. *The Modern Language Journal, 81*(3), 344-362.

Gay, L., Mills, G., & Airasian, P. (2009). *Educational Research* (Vol. (9th ed.)). Pearson plc.

GDA. (2013). *General Directorate of Assessement: The structure of exams in public schools*.

Ghanizadeh, A. (2011). An investigation into the relationship between self-regulation and critical thinking among Iranian EFL teachers. *Technology of Education Journal (TEJ), 5*(2), 117-124.

Ghanizadeh, A. (2017). The interplay between reflective thinking, critical thinking, self-monitoring, and academic achievement in higher education. *Higher Education, 74*(1), 101-114. (Higher Education)

Ghanizadeh, A., & Mirzaee, S. (2012). EFL learners' self-regulation, critical thinking and language achievement. *International Journal of Linguistics, 4*(3), 451-468. (International Journal of Linguistics)

Gialdino, & Vasilachis, I. (2009). Ontological and epistemological foundations of qualitative research. [Forum Qualitative Sozialforschung/Forum: Qualitative Social Research]. Forum Qualitative Sozialforschung/Forum: Qualitative Social Research,

Gibson, W., & Brown, A. (2009). *Working with qualitative data*. Sage.

Gillham, B. (2000). *Case study research methods*. Bloomsbury Publishing.

Gipps, C. (1994). *Beyond Testing: Towards a theory of educational assessment*. Routledge.

Gipps, C. (2002). Sociocultural Perspectives on Assessment. In *Learning for Life in the 21st Century* (pp. 73-83). https://doi.org/https://doi.org/10.1002/9780470753545.ch6

Glassman, M. (2001). Dewey and Vygotsky: Society, experience, and inquiry in educational practice. *Educational researcher, 30*(4), 3-14.

Goh, J. W., Lee, O. K., & Salleh, H. (2010). Self-rating and respondent anonymity. *Educational Research, 52*(3), 229-245.

Golafshani, N. (2003). Understanding reliability and validity in qualitative research. *The qualitative report, 8*(4), 597-607. (The qualitative report)

Goodrich, H. W. (1996). *Student self-assessment: At the intersection of metacognition and authentic assessment*. Harvard University.

Gordon, E. W., & Rajagopalan, K. (2016). The Gordon Commission and a Vision for the Future of Assessment in Education. In *The Testing and Learning Revolution* (pp. 1-8). Springer.

Graham, S., Hebert, M., & Harris, K. R. (2015). Formative assessment and writing: A meta-analysis. *The Elementary School Journal, 115*(4), 523-547.

Griffith, W., & Lim, H. (2012). Performance-based assessment: rubrics, web 2.0 tools and language competencies. *Mextesol Journal, 36*(1).

Griffiths, C. (2007). Language learning strategies: students' and teachers' perceptions. *ELT journal, 61*(2), 91-99.

Griffiths, C., & Parr, J. M. (2001). Language-learning strategies: Theory and perception. *ELT journal, 55*(3), 247-254. (ELT journal)

Guba, E. G., & Lincoln, Y. S. (1994). Competing paradigms in qualitative research. *Handbook of qualitative research, 2*(163-194), 105. (Handbook of qualitative research)

Guettal, I. (2008). *Introducing Storytelling Activities as a Basic Technique to Develop Students' Oral PerformanceAn Experimental Approach Case Study: Second Year Students of English at Batna University* Universite Mohamed Khider].

Guillemin, M., & Gillam, L. (2004). Ethics, reflexivity, and "ethically important moments" in research. *Qualitative inquiry, 10*(2), 261-280. (Qualitative inquiry)

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American psychologist, 53*(4), 449.

Halpern, D. F. (1999). Teaching for critical thinking: Helping college students develop the skills and dispositions of a critical thinker. *New directions for teaching and learning, 1999*(80), 69-74.

Hamby, B. (2014). *The virtues of critical thinkers*

Han, H.-S. (2007). *Family influence on children's second language literacy building: A case study of Korean families* Indiana University of Pennsylvania].

Handley, K., Price, M., & Millar, J. (2011). Beyond 'doing time': investigating the concept of student engagement with feedback. *Oxford Review of Education, 37*(4), 543-560.

Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: Necessary and illuminating.

Hargreaves, G. (2006). An introduction to assessment: Teaching and learning. *London: The Higher Education Academy*.

Harlen, W. (2005). Teachers' summative practices and assessment for learning–tensions and synergies. *Curriculum Journal, 16*(2), 207-223.

Harlen, W., & Deakin Crick, R. (2003). Testing and motivation for learning. *Assessment in Education: principles, policy & practice, 10*(2), 169-207.

Harlen, W., & James, M. (1997). Assessment and learning: differences and relationships between formative and summative assessment. *Assessment in Education: principles, policy & practice, 4*(3), 365-379.

Harris, M. (1997). Self-assessment of language learning in formal settings. *ELT journal, 51*(1), 12-20.

Hashemi, M. R., & Zabihi, R. (2012). Does Critical Thinking Enhance EFL Learners' Receptive Skills? *Journal of Language Teaching & Research, 3*(1). (Journal of Language Teaching & Research)

Hassan, K. E., & Madhum, G. (2007). Validating the Watson Glaser critical thinking appraisal. *Higher Education, 54*(3), 361-383. (Higher Education)

Hattie, J., & Clarke, S. (2018). *Visible learning: feedback*. Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112.

Haughton, G., & Dickinson, L. (1988). Collaborative assessment by masters' candidates in a tutor based system. *Language testing, 5*(2), 233-246.

Hayes, N. (1997). *Doing qualitative analysis in psychology*. Psychology Press.

Hayes, N. (2013). *Doing qualitative analysis in psychology*. Psychology Press.

Hayward, L. (2013). Has anyone really learned anything? In D. Wyse, B. Vivienne, E. David, H. Louise, & G. Carmel (Eds.), *Creating the Curriculum. Series: Understanding primary education series. Routledge: London*.

Healy, M., & Perry, C. (2000). Comprehensive criteria to judge validity and reliability of qualitative research within the realism paradigm. *Qualitative market research: An international journal*. (Qualitative market research: An international journal)

Herayati, D. F. (2020). Teaching Writing Through Self-Assessment And Analytical Scoring.

Herbert, W., Seliger, H. W., Shohamy, E. G., & Shohamy, E. (1989). *Second language research methods*. Oxford University Press.

Herman, J. L. (1992). *A practical guide to alternative assessment*. ERIC.

Heywood, J. (2000). *Assessment in higher education: Student learning, teaching, programmes and institutions* (Vol. 56). Jessica Kingsley Publishers.

Hiep, P. H. (2007). Communicative language teaching: Unity within diversity. *ELT journal, 61*(3), 193-201.

Hitchcock, D. (2018a). *Assessment*. Stanford Encyclopedia of Philosophy.

Hitchcock, D. (2018b). *Critical thinking*. Stanford Encyclopedia of Philosophy. https://stanford.library.sydney.edu.au/archives/spr2019/entries/critical-thinking/

Hitchcock, D. (2018c). *Internal Critical Thinking Dispositions*. Stanford Encyclopedia of Philosophy. https://plato.stanford.edu/entries/critical-thinking/dispositions.html

Hitchcock, G., & Hughes, D. (2002). *Research and the teacher: A qualitative introduction to school-based research*. Routledge.

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and achievement in problem-based and inquiry learning: a response to Kirschner, Sweller, and. *Educational psychologist, 42*(2), 99-107.

Hoepfl, M. C. (1997). Choosing qualitative research: A primer for technology education researchers. *Volume 9 Issue 1 (fall 1997)*. (Volume 9 Issue 1 (fall 1997))

Hofer, B. K. (2008). Personal epistemology and culture. In *Knowing, knowledge and beliefs* (pp. 3-22). Springer.

Holec, H. (1981). *Autonomy and foreign language learning*. ERIC.

Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal, 70*(2), 125-132.

Hughes, A. (2003). *Testing for language teachers*. Cambridge university press.

Hunt, J., Gow, L., & Barnes, P. (1989). Learner self-evaluation and assessment-a tool for autonomy in the language learning classroom. *Language teaching and learning styles within and across cultures*, 207-217.

Hurd, S., & Lewis, T. (2008). *Language learning strategies in independent settings* (Vol. 33). Multilingual matters Bristol.

Immordino-Yang, M. H. (2015). *Emotions, learning, and the brain: Exploring the educational implications of affective neuroscience (the Norton series on the social neuroscience of education)*. WW Norton & Company.

Isaacs, T., & Isaacs, T. (2016). Assessing speaking. *Handbook of second language assessment, 12*, 131-146.

Jafari, Z., Yavari, S., & Ahmadi, S. D. (2015). THE IMPACT OF SELF-ASSESSMENT ON EFL LEARNERS'CRITICAL THINKING. *Modern Journal of Language Teaching Methods, 5*(1), 145.

James, M. (2006). Assessment, teaching and theories of learning. *Assessment and learning, 47*, 60.

Javid, C. Z., Farooq, U., & Gulzar, M. A. (2012). Saudi English-major undergraduates and English Teachers' perceptions regarding effective ELT in the KSA: A Comparative Study. *European Journal of Scientific Research, 85*(1), 55-70. (European Journal of Scientific Research)

Jessop, T., & Maleckar, B. (2016). The influence of disciplinary assessment patterns on student learning: A comparative study. *Studies in Higher Education, 41*(4), 696-711.

Johnson, R. B., & Christensen, L. (2004). *Educational research: Quantitative, qualitative, and mixed approaches*. Sage publications.

Johnson, R. B., & Onwuegbuzie, A. J. (2004). Mixed methods research: A research paradigm whose time has come. *Educational researcher, 33*(7), 14-26. (Educational researcher)

Jones, M. G., Jones, B. D., Hardin, B., Chapman, L., Yarbrough, T., & Davis, M. (1999). The Impact of High-Stakes Testing on Teachers and Students in North Carolina. *The Phi Delta Kappan, 81*(3), 199-203. http://www.jstor.org/stable/20439620

Jupp, V. (2006). *The Sage dictionary of social research methods*. sage.

Kachru, B. B., Kachru, Y., & Nelson, C. L. (2009). *The handbook of world Englishes* (Vol. 48). John Wiley & Sons.

Kahrizi, P., & Farahian, M. (2014). The Relationship Between Self-Regulation and Critical Thinking Among Iranian EFL Learners. *Modern Journal of Language Teaching Methods, 4*(4), 157.

Kahrizi, P., Farahian, M., & Rajabi, S. (2014). The impact of self-assessment on self-regulation and critical thinking of EFL learners. *Modern Journal of Language Teaching Methods, 4*(1), 353.

Kalra, R., Sundrarajun, C., & Komintarachat, H. (2017). Using portfolio as an alternative assessment tool to enhance Thai EFL students' writing skill. *Arab World English Journal (AWEJ) Volume, 8*.

Kamgar, N., & Jadidi, E. (2016). Exploring the Relationship of Iranian EFL Learners□ Critical Thinking and Self-regulation with their Reading Comprehension Ability. *Procedia-Social and Behavioral Sciences, 232*, 776-783. (Procedia-Social and Behavioral Sciences)

Kariri, K., Cobern, W. W., & Bentz, A. (2018). Investigating the Use of Formative Assessment among Male Saudi Arabian High School Science Teachers. *International Journal of Environmental and Science Education, 13*(1), 889-901. (International Journal of Environmental and Science Education)

Kasule, D., & Lunga, V. B. (2010). Attitudes of second language students towards self-editing their own written texts. *Reading & Writing-Journal of the Reading Association of South Africa, 1*(1), 61-72.

Kennedy, K. J., Chan, J. K. S., Fok, P. K., & Yu, W. M. (2008). Forms of assessment and their potential for enhancing learning: Conceptual and cultural issues. *Educational Research for Policy and Practice, 7*(3), 197-207.

Keppell, M., Au, E., Ma, A., & Chan, C. (2006). Peer learning and learning-oriented assessment in technology-enhanced environments. *Assessment & evaluation in higher education, 31*(4), 453-464. (Assessment & Evaluation in Higher Education)

Khan, I. A. (2011). Learning difficulties in English: Diagnosis and pedagogy in Saudi Arabia. *Educational Research, 2*(7), 1248-1257.

Klenowski, V. (1995). Student self-evaluation processes in student-centred teaching and learning contexts of Australia and England. *Assessment in Education: principles, policy & practice, 2*(2), 145-163.

Klenowski, V. (2009). Assessment for learning revisited: An Asia-Pacific perspective.

Köller, O. (2005). Formative assessment in classrooms: A review of the empirical German literature. *Formative assessment: improving learning in secondary classrooms*, 265-279.

Ku, K. Y., & Ho, I. T. (2010a). Dispositional factors predicting Chinese students' critical thinking performance. *Personality and individual differences, 48*(1), 54-58.

Ku, K. Y., & Ho, I. T. (2010b). Metacognitive strategies that enhance critical thinking. *Metacognition and learning, 5*(3), 251-267.

Kumar, R. (2019). *Research methodology: A step-by-step guide for beginners*. Sage.

Kvale, S. (1996). The 1,000-page question. *Qualitative inquiry, 2*(3), 275-284. (Qualitative inquiry)

Kvale, S., & Brinkmann, S. (2009). *Interviews: Learning the craft of qualitative research interviewing*. sage.

Kyndt, E., Dochy, F., Struyven, K., & Cascallar, E. (2011). The perception of workload and task complexity and its influence on students' approaches to learning: A study in higher education. *European journal of psychology of education, 26*(3), 393-415.

Lajoie, S. P. (2008). Metacognition, self regulation, and self-regulated learning: A rose by any other name? *Educational psychology review, 20*(4), 469-475.

Lambert, D., & Lines, D. (2013). *Understanding assessment: Purposes, perceptions, practice*. Routledge.

Lau, A. M. S. (2016, 2016/07/03). 'Formative good, summative bad?' – A review of the dichotomy in assessment literature. *Journal of Further and higher education, 40*(4), 509-525. https://doi.org/10.1080/0309877X.2014.984600

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation*. Cambridge university press.

Lebeeb, R. (1993). The General Principles for Teaching in Beirut *Lebanon, Dar-Aludah*.

Lee, I., & Mak, P. (2014). Assessment as learning in the language classroom. *Assessment as learning. Hong Kong: Education Bureau, 3*, 66-78.

Leont'ev, A. N. (1978). *Activity, consciousness, and personality*.

Levelt, W. J. (1994). The skill of speaking. *International perspectives on psychological science, 1*, 89-103.

Lillis, A. (2006). Reliability and validity in field study research. *Methodological issues in accounting research: Theories and methods*, 461-475.

Lincoln, Y. S., & Guba, E. G. (1985). *Naturalistic inquiry*. sage.

Lincoln, Y. S., & Guba, E. G. (2003). Ethics: The failure of positivist science. *Turning points in qualitative research: Tying knots in a handkerchief*, 219-238.

Lindblom-Ylänne, S., Trigwell, K., Nevgi, A., & Ashwin, P. (2006). How approaches to teaching are affected by discipline and teaching context. *Studies in Higher Education, 31*(03), 285-298.

Linnenbrink, E. A., & Pintrich, P. R. (2002). Motivation as an enabler for academic success. *School psychology review, 31*(3), 313-327.

Little, D. G. (1991). *Learner autonomy: Definitions, issues and problems*. Authentik Language Learning Resources.

Liu, H., & Brantmeier, C. (2019). "I know English": Self-assessment of foreign language reading and writing abilities among young Chinese learners of English. *System, 80*, 60-72.

Liu, M., & Huang, W. (2011). An exploration of foreign language anxiety and English learning motivation. *Education Research International, 2011*.

Livingston, K., & Hutchinson, C. (2017). Developing teachers' capacities in assessment through career-long professional learning. *Assessment in Education: principles, policy & practice, 24*(2), 290-307.

Lowe, K., & Dotterer, A. M. (2018). Parental involvement during the college transition: A review and suggestion for its conceptual definition. *Adolescent Research Review, 3*(1), 29-42.

Lucieer, S. M., Jonker, L., Visscher, C., Rikers, R. M., & Themmen, A. P. (2016). Self-regulated learning and academic performance in medical education. *Medical Teacher, 38*(6), 585-593. (Medical teacher)

Lukka, K. (1988). Budgetary biasing in organizations: theoretical framework and empirical evidence. *Accounting, organizations and society, 13*(3), 281-301. (Accounting, organizations and society)

Luoma, S. (2004). *Assessing speaking*. Cambridge University Press.

Maehr, M. L., & Zusho, A. (2009). Achievement Goal Th eory: Th e Past, Present, and Future. In *Handbook of motivation at school* (pp. 91-118). Routledge.

Mahmoud, S. (2012). The effect of using L1 (Arabic language) in the L2 (English Language) classroom on the achievement in general English of foundation year students in King Abdulaziz University. *Sino-US English Teaching, 9*(12), 1733-1738.

Manso-Vázquez, M., Caeiro-Rodríguez, M., & Llamas-Nistal, M. (2014). A questionnaire for the evaluation of self-regulated learning support in software. 2014 International Symposium on Computers in Education (SIIE),

Manso-Vázquez, M., Caeiro-Rodríguez, M., & Llamas-Nistal, M. (2018). An xAPI application profile to monitor self-regulated learning strategies. *IEEE Access, 6*, 42467-42481. (IEEE Access)

Marks, D. F., & Yardley, L. (2004). *Research methods for clinical and health psychology*. Sage.

Marshall, B., & Drummond, M. (2006). How teachers engage with assessment for learning: Lessons from the classroom. *Research papers in education, 21*(02), 133-149.

Maryam, A., Mohammadreza, D., Abdolhussein, S., Ghobad, R., & Javad, K. (2021). Effect of Concept Mapping Education on Critical Thinking Skills of Medical Students: A Quasi-experimental Study. *Ethiopian journal of health sciences, 31*(2). (Ethiopian journal of health sciences)

Mason, L., Boldrin, A., & Ariasi, N. (2010). Epistemic metacognition in context: evaluating and learning online information. *Metacognition and learning, 5*(1), 67-90.

Massri, R. (2017). *Attitudes of Saudi foundation year students towards learning English as a Foreign Language: A qualitative study* University of York].

McCarthy, P., Meier, S., & Rinderer, R. (1985). Self-Efficacy and Writing: A Different View of Self-Evaluation. *College Composition and Communication, 36*(4), 465-471. https://doi.org/10.2307/357865

McChesney, K., & Aldridge, J. (2019). Weaving an interpretivist stance throughout mixed methods research. *International journal of research & method in education, 42*(3), 225-238. (International journal of research & method in education)

McLeod, S. (1997). Student self-appraisal: Facilitating mutual planning in clinical education. *The Clinical Supervisor, 15*(1), 87-101.

McMillan, J. H., & Hearn, J. (2008). Student self-assessment: The key to stronger student motivation and higher achievement. *Educational horizons, 87*(1), 40-49. (Educational horizons)

Merriam, S. B., & Tisdell, E. J. (2015). *Qualitative research: A guide to design and implementation*. John Wiley & Sons.

MoE. (1970). *Educational Policy in the Kingdom of Saudi Arabia.* Riyadh, Saudi Arabia: The Ministry of Education.

MoE, T. M. o. E. (2005). *Goals and Objectives of Teaching English in Saudi Arabia.*

Mohammed, A. (2016). EFL effective factors: anxiety and motivation and their effect on Saudi college student's achievement. *Arab World English Journal (AWEJ) Vol, 6*.

Mok, M. M. C. (2012). Assessment Reform in the Asia-Pacific Region: The theory and practice of self-directed learning oriented assessment. In *Self-directed learning oriented assessments in the Asia-Pacific* (pp. 3-22). Springer.

Montgomery, W. (2001). Creating culturally responsive, inclusive classrooms. *Teaching Exceptional Children, 33*(4), 4-9.

Morrow, S. L. (2005). Quality and trustworthiness in qualitative research in counseling psychology. *Journal of counseling psychology, 52*(2), 250. (Journal of counseling psychology)

Mousavi, S. A. (1999). *A dictionary of language testing*. Rahmana Publications.

Muis, K. R. (2007). The role of epistemic beliefs in self-regulated learning. *Educational psychologist, 42*(3), 173-190.

Muis, K. R., Chevrier, M., & Singh, C. A. (2018). The role of epistemic emotions in personal epistemology and self-regulated learning. *Educational psychologist, 53*(3), 165-184.

Munns, G., & Woodward, H. (2006). Student engagement and student self-assessment: the REAL framework. *Assessment in Education: principles, policy & practice, 13*(2), 193-213.

Murawski, L. M. (2014). Critical Thinking in the Classroom… and Beyond. *Journal of Learning in Higher Education, 10*(1), 25-30.

Murphy, L. (2008). Supporting learner autonomy: Developing practice through the production of courses for distance learners of French, German and Spanish. *Language teaching research, 12*(1), 83-102.

Murphy, P., Sharp, G., & Whitelegg, E. (2006). Girls' experience of physics: A problem of identification and marginalization. *Unpublished journal article. The Open University*.

Murray, G. (2014). Exploring the social dimensions of autonomy in language learning. In *Social dimensions of autonomy in language learning* (pp. 3-11). Springer.

Murtagh, L., & Webster, M. (2010). Scaffolding teaching, learning and assessment in Higher Education. *Teacher Education Advancement Network Journal, 1*(2).

Naber, J., & Wyatt, T. H. (2014). The effect of reflective writing interventions on the critical thinking skills and dispositions of baccalaureate nursing students. *Nurse Education Today, 34*(1), 67-72. (Nurse Education Today)

Nalliveettil, G. M., & Mahasneh, A. (2017). Developing competence in basic writing skills: perceptions of EFL undergraduates. *International Journal of Applied Linguistics and English Literature, 6*(7), 323-341.

Natriello, G. (1987). The impact of evaluation processes on students. In *Educational Psychologist,* (pp. 155–175). Routledge.

NCAAA. (2015). *Handbook for quality assurance and accreditation in Saudi Arabia part 2: Internal quality assurance arrangements*. (NCAAA)

Nelson, K. L., Rauter, C. M., & Cutucache, C. E. (2018). Life Science undergraduate mentors in NE STEM 4U significantly outperform their peers in critical thinking

skills. *CBE—Life Sciences Education, 17*(4), ar54. (CBE—Life Sciences Education)

Neuman, W. L., & Kreuger, L. (2003). *Social work research methods: Qualitative and quantitative approaches*. Allyn and Bacon.

Neumann, R. (2001). Disciplinary differences and university teaching. *Studies in Higher Education, 26*(2), 135-146.

Neumann, R., Parry, S., & Becher, T. (2002). Teaching and learning in their disciplinary contexts: A conceptual analysis. *Studies in Higher Education, 27*(4), 405-417.

Newton, P. E. (2007). Clarifying the purposes of educational assessment. *Assessment in education, 14*(2), 149-170.

Newton, P. E. (2010). The multiple purposes of assessment. In *International Encyclopaedia of Education.* (pp. 392-396).

Nicol, D. (2010). From monologue to dialogue: improving written feedback processes in mass higher education. *Assessment & evaluation in higher education, 35*(5), 501-517.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Nielsen, K. (2011). *Peer evaluation and self assessment: A comparative study of the effectiveness of two complex methods of writing instruction in six sections of composition* Boston University].

Nonneman, G. (2001). Saudi—European relations 1902–2001: a pragmatic quest for relative autonomy. *International Affairs, 77*(3), 631-661.

Nouraldeen, A. S., & Elyas, T. (2014). Learning English in Saudi Arabia: A socio-cultural perspective. *International Journal of English Language and Linguistics Research, 2*(3), 56-78. (International Journal of English Language and Linguistics Research)

Nunan, D. (1989). A client-centred approach to teacher development. *ELT journal, 43*(2), 111-118.

Nunan, D., David, N., & Swan, M. (1992). *Research methods in language learning*. Cambridge university press.

O'malley, J., & Pierce, L. (1996). Authentic Assessment For English Language Learners: Practical Approaches For Teachers Author: J. *Michael O'Malley, Lorrain*.

Obias, P. H. (2015). Critical Thinking of College Students: Inputs to Teacher Education Curriculum. *The Normal Lights, 9*(2). (The Normal Lights)

Onwuegbuzie, A. J., & Johnson, R. B. (2006). The validity issue in mixed research. *Research in the Schools, 13*(1), 48-63. (Research in the Schools)

Opp-Beckman, L., & Klinghammer, S. (2006). Shaping the Way We Teach English: Successful Practices around the World. Readings and Resources. *Washington, DC*.

Oppong, E., Shore, B. M., & Muis, K. R. (2019). Clarifying the connections among giftedness, metacognition, self-regulation, and self-regulated learning: Implications for theory and practice. *Gifted Child Quarterly, 63*(2), 102-119.

Oscarson, A. D. (2009). *Self-Assessment of Writing in Learning English as a Foreign Language: A Study at the Upper Secondary School Level. Goteborg Studies in Educational Sciences 277*. ERIC.

Oscarson, M. (1989). Self-assessment of language proficiency: Rationale and applications. *Language testing, 6*(1), 1-13.

Oskarsson, M. (1984). *Self-Assessment of Foreign Language Skills: A Survey of Research and Development Work*. ERIC.

Oyaid, A. (2009). Education policy in Saudi Arabia and its relation to secondary school teachers' ICT use, perceptions, and views of the future of ICT in education.

Pallant, J. (2010). A step by step guide to data analysis using the SPSS program: SPSS survival manual. *England: McGraw Hill Education and Open University Press. ISBN, 13*, 978-970.

Pallant, J. (2016). *Survival manual: A step by step guide to data analysis using SPSS program*. London: McGraw-Hill Education.

Palys, T. (2008). Purposive sampling. *The Sage encyclopedia of qualitative research methods, 2*(1), 697-698.

Panadero, Alonso-Tapia, J., & Reche, E. (2013). Rubrics vs. self-assessment scripts effect on self-regulation, performance and self-efficacy in pre-service teachers. *Studies in Educational Evaluation, 39*(3), 125-132. https://doi.org/10.1016/j.stueduc.2013.04.001

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology, 8*, 422-422. https://doi.org/10.3389/fpsyg.2017.00422

Panadero, E., & Alonso-Tapia, J. (2013). Self-assessment: theoretical and practical connotations, when it happens, how is it acquired and what to do to develop it in our students.

Panadero, E., Andrade, H., & Brookhart, S. (2018). Fusing self-regulated learning and formative assessment: A roadmap of where we are, how we got here, and where we are going. *The Australian Educational Researcher, 45*(1), 13-31.

Panadero, E., Brown, G., & Courtney, M. (2013). What teachers think about and how they use self-assessment: An approach to its use in Spain. *Sent for publication*.

Panadero, E., Brown, G., & Strijbos, J.-W. (2016). The future of student self-assessment: A review of known unknowns and potential directions. *Educational psychology review, 28*(4), 803-830.

Panadero, E., & Jonsson, A. (2013). The use of scoring rubrics for formative assessment purposes revisited: A review. *Educational Research Review, 9*, 129-144.

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review, 22*, 74-98.

Panadero, E., Jonsson, A., & Strijbos, J.-W. (2016). Scaffolding self-regulated learning through self-assessment and peer assessment: Guidelines for classroom implementation. Assessment for learning: Meeting the challenge of implementation,

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.

Panadero, E., Tapia, J. A., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and individual differences, 22*(6), 806-813.

Papanthymou, A., & Darra, M. (2018). Student Self-Assessment in Higher Education: The International Experience and The Greek Example. *World journal of education, 8*(6), 130-146.

Patchen, T., & Smithenry, D. W. (2014). Diversifying instruction and shifting authority: A cultural historical activity theory (CHAT) analysis of classroom participant structures. *Journal of Research in Science Teaching, 51*(5), 606-634.

Patton, M. Q. (1990). *Qualitative evaluation and research methods*. SAGE Publications, inc.

Patton, M. Q. (2015). Enhancing the quality and credibility of qualitative studies. *Qualitative research & evaluation methods*, 652-743.

Paul, R., & Elder, L. (2006). Critical thinking: The nature of critical and creative thought. *Journal of Developmental Education, 30*(2), 34.

Paul, R. W., Elder, L., & Bartell, T. (1997). California teacher preparation for instruction in critical thinking: Research findings and policy recommendations.

Pemberton, R., Li, E. S., Or, W. W., & Pierson, H. D. (1996). *Taking control: Autonomy in language learning* (Vol. 1). Hong kong University press.

Perry, N. E. (1998). Young children's self-regulated learning and contexts that support it. *Journal of educational psychology, 90*(4), 715. (Journal of educational psychology)

Perry, N. E., & Winne, P. H. (2006). Learning from learning kits: gStudy traces of students' self-regulated engagements with computerized content. *Educational psychology review, 18*(3), 211-228. (Educational Psychology Review)

Phan, H. P. (2010). Critical thinking as a self-regulatory process component in teaching and learning. *Psicothema*, 284-292.

Piaget, J. (1978). *The development of thought: Equilibration of cognitive structures.(Trans A. Rosin)*. Viking.

Picciano, A. (2004). *Educational research primer*. A&C Black.

Piccinin, S. J. (2003). *Feedback: key to learning*. Halifax, NS: Society for Teaching and Learning in Higher
Education.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In *Handbook of self-regulation* (pp. 451-502). Elsevier.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational psychology review, 16*(4), 385-407.

Pintrich, P. R., Conley, A. M., & Kempler, T. M. (2003). Current issues in achievement goal theory and research. *International Journal of Educational Research, 39*(4-5), 319-337.

Pitt, E., & Winstone, N. (2020). Towards technology enhanced dialogic feedback. In *Re-imagining university assessment in a digital world* (pp. 79-94). Springer.

Powell, W. W., & Snellman, K. (2004). The knowledge economy. *Annual review of sociology*, 199-220.

Price, M., Handley, K., & Millar, J. (2011). Feedback: Focusing attention on engagement. *Studies in Higher Education, 36*(8), 879-896.

Punch, K. F. (2009). *Introduction to Research Methods in Education*. SAGE.

Purpura, J., & Turner, C. (2014). A learning-oriented assessment approach to understanding the complexities of classroom-based language assessment. *Roundtable on learning-oriented Assessment in Language Classrooms and Large-Scale Contexts*.

Putri, N. V. W., & Munir, A. (2021). Students' perceptions of teacher feedback in EFL English class and their self-regulated learning after receiving feedback. *Journal on English as a Foreign Language, 11*(1), 42-60.

Qasem, F. A. A. (2020). The Effective Role of Learners' Self-Assessment Tasks in Enhancing Learning English as a Second Language. *Arab World English Journal, 11*(3), 502-514.

Rabiah, M. N. A. B. (2020). Effectiveness of Self-Assessment Technique in Reading Skills of EFL Students. *Multi-knowledge Electronic Comprehensive Journal Foe Education And Science Publications (MECSJ)*(36), 1-27.

Ramaprasad, A. (1983). On the definition of feedback. *Behavioral science, 28*(1), 4-13.

Ratminingsih, N. M., Marhaeni, A., & Vigayanti, L. (2018). Self-Assessment: The Effect on Students' Independence and Writing Competence. *International Journal of Instruction, 11*(3), 277-290.

Ratner, C. (2000). Agency and culture. *Journal for the theory of social behaviour*.

Reitano, P. (2006). The value of video stimulated recall in reflective teaching practices. Proceedings of the Australian Consortium for Social and Political Research Incorporated (ACSPRI) Social Sciences Methodology Conference,

Rheinberg, F., Vollmeyer, R., & Rollett, W. (2000). Motivation and action in self-regulated learning. In *Handbook of self-regulation* (pp. 503-529). Elsevier.

Richards, J. C. (2006). *Communicative language teaching today*. SEAMEO Regional Language Centre Singapore.

Richards, J. C., Richards, J. C., & Renandya, W. A. (2002). *Methodology in language teaching: An anthology of current practice*. Cambridge university press.

Richards, J. C., & Rodgers, T. S. (2001). *Approaches and methods in language teaching*. Cambridge University Press.

Richards, J. C., & Rodgers, T. S. (2014). *Approaches and methods in language teaching*. Cambridge university press.

Rivers, W. P. (2001). Autonomy at all costs: An ethnography of metacognitive self-assessment and self-management among experienced language learners. *The Modern Language Journal, 85*(2), 279-290. (The modern language journal)

Roghanizadeh, L. (2011). An investigation into the relationship between self-regulation and critical thinking among Iranian EFL teachers.

Ross, J. A. (2006). The reliability, validity, and utility of self-assessment. *Practical Assessment, Research, and Evaluation, 11*(1), 10.

Rowe, F. A., & Rafferty, J. A. (2013). Instructional design interventions for supporting self-regulated learning: enhancing academic outcomes in postsecondary e-learning environments. *Journal of Online Learning and Teaching, 9*(4), 590-601.

Rychtarik, S. (2014). *Designing criteria to assess speaking skills* Doctoral Dissertation. From< https://www/. academia. edu/7719276/Designing …].

Saad, S., Carter, G. W., Rothenberg, M., & Israelson, E. (1999). Testing and Assessment: An Employer's Guide to Good Practices.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional science, 18*(2), 119-144.

Sadler, D. R. (2010). Beyond feedback: Developing student capability in complex appraisal. *Assessment & evaluation in higher education, 35*(5), 535-550.

Saeed, S., & Zyngier, D. (2012). How motivation influences student engagement: A qualitative case study. *Journal of Education and learning, 1*(2), 252-267.

Saks, K. (2016). *Supporting students' self-regulation and language learning strategies in the blended course of Professional English* Tartu University].

Salkind, N. J. (2010). *Encyclopedia of research design* (Vol. 1). sage.

Sambell, K., McDowell, L., & Montgomery, C. (2012). *Assessment for learning in higher education*. Routledge.

Sanchez, C. E., Atkinson, K. M., Koenka, A. C., Moshontz, H., & Cooper, H. (2017). Self-grading and peer-grading for formative and summative assessments in 3rd through 12th grade classrooms: A meta-analysis. *Journal of educational psychology, 109*(8), 1049.

Sandhu, B., & Sharma, S. (2015). Role of critical thinking in ego identity statuses among adolescents. *Indian Journal of Health & Wellbeing, 6*(11). (Indian Journal of Health & Wellbeing)

Saudi-Vision-2030. (2016). *Kingdom of Saudi Arabia: Vision 2030*. http://vision2030.gov.sa/en

Saunders, M., Lewis, P., & Thornhill, A. (2019). *Research methods for business students*. Pearson education.

Sax, L. J., & Wartman, K. L. (2010). Studying the impact of parental involvement on college student development: A review and agenda for research. *Higher education: Handbook of theory and research*, 219-255.

Scheffler, I. (1960). The language of education. *Philosophy, 38*(144), 19.

Schmidt, R. (1992). Psychological mechanisms underlying second language fluency. *Studies in second language acquisition, 14*(4), 357-385.

Scholten, L., Van Knippenberg, D., Nijstad, B. A., & De Dreu, C. K. (2007). Motivated information processing and group decision-making: Effects of process accountability on information processing and decision quality. *Journal of Experimental Social Psychology, 43*(4), 539-552.

Schunk, D. H. (2001). Self-regulation through goal setting.

Schunk, D. H., & Zimmerman, B. J. (2012). *Motivation and self-regulated learning: Theory, research, and applications*. Routledge.

Scotland, J. (2012). Exploring the philosophical underpinnings of research: Relating ontology and epistemology to the methodology and methods of the scientific, interpretive, and critical research paradigms. *English Language Teaching, 5*(9), 9-16. (English language teaching)

Scriven, M. (1967). *The methodology of evaluation. Perspectives of Curriculum Evaluation*. Washington, DC: American Educational Research Association.

Senko, C., Hulleman, C. S., & Harackiewicz, J. M. (2011). Achievement goal theory at the crossroads: Old controversies, current challenges, and new directions. *Educational psychologist, 46*(1), 26-47.

Shah, S. R., & Al-Bargi, A. (2013). Research Paradigms: Researchers' Worldviews, Theoretical Frameworks and Study Designs. *Arab World English Journal, 4*(4). (Arab World English Journal)

Shah, S. R., Hussain, M. A., & Nasseef, O. A. (2013). Factors Impacting EFL Teaching: An Exploratory Study in the Saudi Arabian Context. *Arab World English Journal, 4*(3).

Sierens, E., Vansteenkiste, M., Goossens, L., Soenens, B., & Dochy, F. (2009). The synergistic relationship of perceived autonomy support and structure in the prediction of self-regulated learning. *British Journal of Educational Psychology, 79*(1), 57-68.

Simonson, M., Smaldino, S., Albright, M., & Zvacek, S. (2000). Teaching and Learning at a Distance: Foundations of Distance Education.

Sitzmann, T., Ely, K., Brown, K. G., & Bauer, K. N. (2010). Self-assessment of knowledge: A cognitive learning or affective measure? *Academy of Management Learning & Education, 9*(2), 169-191.

Sluijsmans, D., Dochy, F., & Moerkerke, G. (1998). Creating a learning environment by using self-, peer-and co-assessment. *Learning environments research, 1*(3), 293-319. (Learning environments research)

Smeby, J.-C. (1996). Disciplinary differences in university teaching. *Studies in Higher Education, 21*(1), 69-79.

Smith, C. (1997). Student self-assessment at St. Bernadette's primary school. *Primary Educator, 3*(4), 7.

Smith, J., & Osborn, M. (2003). Interpretative Phenomenological Analysis. *Qualitative Psychology, 7(2)*, 125–130.

Smith, J., & Shinebourne, P. (2012). Interpretative phenomenological analysis. American Psychological Association.

Smith, L., & Abouammoh, A. (2013). Higher education in Saudi Arabia: Reforms, challenges and priorities. In *Higher Education in Saudi Arabia* (pp. 1-12). Springer. (Higher Education in Saudi Arabia)

Sobral, D. T. (2004). Medical students' self-appraisal of first-year learning outcomes: use of the course valuing inventory. *Medical Teacher, 26*(3), 234-238.

Stavros, C., & Westberg, K. (2009). Using triangulation and multiple case studies to advance relationship marketing theory. *Qualitative market research: An international journal*. (Qualitative Market Research: An International Journal)

Sung, Y.-T., Chang, K.-E., Chiou, S.-K., & Hou, H.-T. (2005). The design and application of a web-based self-and peer-assessment system. *Computers & Education, 45*(2), 187-202. (Computers & Education)

Taheri, M. M. H., Ryasi, H. R., Afshar, M., & Mofatteh, M. R. (2014). Comparison between student rating, faculty self-rating and evaluation of faculty members by

heads of respective academic departments in the school of medicine in Birjand University of Medical Sciences in Iran. *Journal of Education and Health Promotion, 3*.

Taras, M. (2005). Assessment–summative and formative–some theoretical reflections. *British journal of educational studies, 53*(4), 466-478.

Taras, M. (2007). Assessment for learning: Understanding theory to improve practice. *Journal of Further and higher education, 31*(4), 363-371.

Taras, M. (2010). Student self-assessment: Processes and consequences. *Teaching in higher education, 15*(2), 199-209.

Taylor, C. (1994). Assessment for measurement or standards: The peril and promise of large-scale assessment reform. *American educational research journal, 31*(2), 231-262.

Toering, T., Elferink-Gemser, M. T., Jonker, L., van Heuvelen, M. J., & Visscher, C. (2012). Measuring self-regulation in a learning context: Reliability and validity of the Self-Regulation of Learning Self-Report Scale (SRL-SRS). *International Journal of Sport and Exercise Psychology, 10*(1), 24-38. (International Journal of Sport and Exercise Psychology)

Tong, S. Y. A. (2011). Exploring students' perception of and reaction to feedback in school-based assessment.

Torrance, H. (2012). Formative assessment at the crossroads: Conformative, deformative and transformative assessment. *Oxford Review of Education, 38*(3), 323-342.

Trochim, W. M. (2005). *Research methods: The concise knowledge base*. Atomic Dog Publishing.

Turri, J., Alfano, M., & Greco, J. (2017). The Stanford Encyclopedia of Philosophy.

Underhill, N. (1991). *Testing spoken language: A handbook of oral testing techniques*. Cambridge University Press.

UNDP/RBAS, U. N. D. P. R. B. f. A. S. (2006). *Quality Assessment of Education Programmes in Arab Universities Regional Overview Report*.

UofG, U. o. G. (2018). *The Guidelines of the Ethics Committee.*

UoS, U. o. S. (2014). *Univariate Analysis*.

Uyar, R. O., Genc, M. M. Y., & Yasar, M. (2018). Prospective preschool teachers' academic achievements depending on their goal orientations, critical thinking dispositions and self-regulation skills. *European Journal of Educational Research, 7*(3), 601-613.

Van Praag, L., Demanet, J., Stevens, P. A., & Van Houtte, M. (2017). 'Everyone has their own qualities': tracking and academic self-appraisal in Flemish secondary education. *Social Psychology of Education, 20*(3), 601-618.

VanPatten, B., & Williams, J. (2015). *Theories in Second Language Acquisition: An Introduction*. Routledge. https://books.google.co.uk/books?id=d-zHoAEACAAJ

Vygotsky, L. S., & Cole, M. (1978). *Mind in society: Development of higher psychological processes*. Harvard university press.

Wang, W. (2017). Using rubrics in student self-assessment: student perceptions in the English as a foreign language writing context. *Assessment & evaluation in higher education, 42*(8), 1280-1292.

Warschauer, M., & Kern, R. (2000). *Network-based language teaching: Concepts and practice*. Cambridge university press.

Watson, G., & Glaser, E. M. (2010). Watson-GlaserTM II critical thinking appraisal. *Technical Manual and User's Guide*. (Technical Manual and User's Guide)

Watson, G. B. G., E. M. (2002). *Watson–Glaser critical thinking appraisal.* .

Webster, T. L., & Willett, G. M. (2019). Educational Intervention Aimed at Teaching Critical Thinking: A Mixed Methods Investigation. *Radiologic Science & Education, 24*(1). (Radiologic Science & Education)

Weiss, K. (2018). Student Self-Assessment Re-Assessed. *Journal of Academic Writing, 8*(2), 161-175.

Wells, G. E., & Claxton, G. E. (2002). *Learning for life in the 21st century: Sociocultural perspectives on the future of education*. Blackwell Publishing.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3-14.

Willey, K., & Gardner, A. (2009). Investigating the potential of self and peer assessment to develop learning oriented assessment tasks. [Annual Conference of Australasian Association for Engineering Education]. Annual Conference of Australasian Association for Engineering Education,

Williamson, S. N. (2007). Development of a self-rating scale of self-directed learning. *Nurse researcher, 14*(2).

Willis, J. (2009). Assessment for learning: a sociocultural approach. Proceedings of the Australian Association for Research in Education International Education Research Conference: Changing Climates: Education for Sustainable Futures,

Wimmer, R., & Dominick, J. (2013). Mass media research. Cengage learning. *Yang-Wallentin, F., Schmidt, P., Davidov, E., & Bamberg, S.(2004). Is there any interaction effect between intention and perceived behavioral control. Methods of Psychological Research Online, 8*(2), 127-157. (Yang-Wallentin, F., Schmidt, P., Davidov, E., & Bamberg, S.(2004). Is there any interaction effect between intention and perceived behavioral control. Methods of Psychological Research Online)

Winne, P., & Hadwin, A. (1998). Studying as self-regulated learning. . In *Metacognition in educational theory and practice* (pp. 277-304).

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27*(4), 551-572. (Contemporary Educational Psychology)

Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In *Handbook of self-regulation* (pp. 531-566). Elsevier.

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting learners' agentic engagement with feedback: A systematic review and a taxonomy of recipience processes. *Educational psychologist, 52*(1), 17-37.

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*, 3087.

Wragg, E. C. (2003). *Assessment and learning in the secondary school*. Routledge.

Wu, X. V., Heng, M. A., & Wang, W. (2015). Nursing students' experiences with the use of authentic assessment rubric and case approach in the clinical laboratories. *Nurse Education Today, 35*(4), 549-555.

Wyse, D., Hayward, L., & Pandya, J. (2015). *The SAGE handbook of curriculum, pedagogy and assessment*. Sage.

Xu, Y. (2019). Scaffolding students' self-assessment of their English essays with annotated samples: A mixed-methods study. *Chinese Journal of Applied Linguistics, 42*(4), 503-526.

Yan, Z. (2020). Self-assessment in the process of self-regulated learning and its relationship with academic achievement. *Assessment & evaluation in higher education, 45*(2), 224-238.

Yan, Z., & Cheng, E. C. K. (2015). Primary teachers' attitudes, intentions and practices regarding formative assessment. *Teaching and Teacher Education, 45*, 128-136. (Teaching and Teacher Education)

Yanow, D., & Schwartz-Shea, P. (2015). *Interpretation and method: Empirical research methods and the interpretive turn*. Routledge.

Yin, R. K. (2009). *Case study research: Design and methods* (Vol. 5). sage.

Zagzebski, L. T., & Zagaebski, L. T. (1996). *Virtues of the mind: An inquiry into the nature of virtue and the ethical foundations of knowledge*. Cambridge University Press.

Zaid, M. A. (1993). *Comprehensive analysis of the current system of teaching English as a foreign language in the Saudi Arabian intermediate schools*. University of Colorado at Boulder.

Zarei, A. A., Esfandiari, R., & Akbari, A. (2016). Self-regulated learning strategies as predictors of reading comprehension. *Journal of Modern Research in English Language Studies, 3*(2), 34-21.

Zascavage, V., Masten, W. G., Schroeder-Steward, J., & Nichols, C. (2007). Comparison of Critical Thinking in Undergraduates and Graduates in Special Education. *International Journal of Special Education, 22*(1), 25-31.

Zhang, Y. (2009). Reading to speak: Integrating oral communication skills. English Teaching Forum, 47 (1), 32-34.

Zimmerman, B., & Schunck, D. (2001). *Self-Regulated Learning & Academic Achievement: Theoretical Perspectives Mahwah, NJ: Lawrence Elbaum*. Associates Publisher.

Zimmerman, B. J. (1989). A social cognitive view of self-regulated academic learning. *Journal of educational psychology, 81*(3), 329.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational psychologist, 25*(1), 3-17.

Zimmerman, B. J. (1995). Self-efficacy and educational development. *Self-efficacy in changing societies, 1*(1), 202-231.

Zimmerman, B. J. (2000a). Attaining self-regulation: A social cognitive perspective. In *Handbook of self-regulation* (pp. 13-39). Elsevier.

Zimmerman, B. J. (2000b). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82-91.

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into practice, 41*(2), 64-70.

Zimmerman, B. J. (2008). Investigating self-regulation and motivation: Historical background, methodological developments, and future prospects. *American educational research journal, 45*(1), 166-183.

Zimmerman, B. J., & Bandura, A. (1994). Impact of self-regulatory influences on writing course attainment. *American educational research journal, 31*(4), 845-862.

Zimmerman, B. J., & Cleary, T. J. (2009). Motives to self-regulate learning: A social cognitive account. In *Handbook of motivation at school* (pp. 261-278). Routledge.

Zimmerman, B. J., & Kitsantas, A. (2005). Homework practices and academic achievement: The mediating role of self-efficacy and perceived responsibility beliefs. *Contemporary Educational Psychology, 30*(4), 397-417.

APPENDICES

Appendix A: Descriptors of the common European framework of reference for languages (CEFR)

Proficient User	C2	Can understand with ease virtually everything heard or read. Can summarise information from different spoken and written sources, reconstructing arguments and accounts in a coherent presentation. Can express him/herself spontaneously, very fluently and precisely, differentiating finer shades of meaning even in more complex situations.
	C1	Can understand a wide range of demanding, longer texts, and recognise implicit meaning. Can express him/herself fluently and spontaneously without much obvious searching for expressions. Can use language flexibly and effectively for social, academic and professional purposes. Can produce clear, well-structured, detailed text on complex subjects, showing controlled use of organisational patterns, connectors and cohesive devices.
Independent User	B2	Can understand the main ideas of complex text on both concrete and abstract topics, including technical discussions in his/her field of specialisation. Can interact with a degree of fluency and spontaneity that makes regular interaction with native speakers quite possible without strain for either party. Can produce clear, detailed text on a wide range of subjects and explain a viewpoint on a topical issue giving the advantages and disadvantages of various options.
	B1	Can understand the main points of clear standard input on familiar matters regularly encountered in work, school, leisure, etc. Can deal with most situations likely to arise whilst travelling in an area where the language is spoken. Can produce simple connected text on topics which are familiar or of personal interest. Can describe experiences and events, dreams, hopes and ambitions and briefly give reasons and explanations for opinions and plans.
Basic User	A2	Can understand sentences and frequently used expressions related to areas of most immediate relevance (e.g. very basic personal and family information, shopping, local geography, employment). Can communicate in simple and routine tasks requiring a simple and direct exchange of information on familiar and routine matters. Can describe in simple terms aspects of his/her background, immediate environment and matters in areas of immediate need.
	A1	Can understand and use familiar everyday expressions and very basic phrases aimed at the satisfaction of needs of a concrete type. Can introduce him/herself and others and can ask and answer questions about personal details such as where he/she lives, people he/she knows and things he/she has. Can interact in a simple way provided the other person talks slowly and clearly and is prepared to help.

Appendix B: A sample of the pacing guides

WEEK 1 - Recruitment-The Gold of Business

STUDENTS LEARNING OUTCOMES				
Theme	Employment and Business			
Thinking Skills	✓ Making predictions ✓ Summarizing			
Study Skills	✓ Skimming and scanning. ✓ Brainstorming			
Presentation Skills	Using proper voice tone. Using proper posture.			
Vocabulary list	1. Voluntary work 2. Unemployed 3. Relocation 4. Technique 5. Assumption 6. Circumstance 7. Recruitment	8. Advertisement 9. Practical skill 10. Profession 11. Applicant 12. Organization 13. Management 14. Influential	15. Candidate 16. Frustrating 17. Flexible 18. Suggest 19. Situation 20. Interview	21. Workplace 22. Attracting 23. Strategy 24. Role 25. Industry

Main SLOs

By the end of this week, students can

- ✓ Make predictions about a content.
- ✓ Connect prior knowledge to their predictions.
- ✓ Take notes to Identify and analyze key information related to "Recruitment and Hiring Human Resources"
- ✓ Develop a topic for Assignment 1
- ✓ Use target vocabulary and grammar forms in meaningful writing and speaking contexts.
- ✓ Identify and use patterns of "modals of deduction in present" in meaningful sentences and contexts.

Skill	Specific SLOs
GRAMMAR & VOCABULARY	✓ Identify and use patterns of "modals of deduction in present" in meaningful sentences and contexts. ✓ Use listed vocabulary in meaningful contexts and sentences
LISTENING	✓ Make predictions about a listening text related to "Recruitment and Hiring Human Resources". ✓ Take notes to Identify and analyze key information related to "Recruitment and Hiring Human Resources" ✓ Identify the basic genre or type of a listening text.

READING	✓ Make predictions of reading text on "Recruitment and Hiring Human Resources" through skimming and based on prior knowledge, visuals and other sources. ✓ Identify and analyze key information related to "Recruitment and hiring human resources" through scanning.
WRITING	✓ Use target weekly vocabulary and grammar forms in writing. ✓ Demonstrate understanding of plagiarism and expectations for academic honesty and how it relates to their work in this course. ✓ Demonstrate understanding of the purpose and basic genres of academic writing. ✓ Demonstrate understanding of the purpose and structure of narrative and descriptive paragraphs. ✓ Choose and develop an appropriate topic for a descriptive or narrative essay of one paragraph, 150 words. ✓ Explore brainstorming as a prewriting strategy. **Suggested Procedures:** Instructor gives students instructions and guidelines for Assignment 1 :Narrative Essay
SPEAKING	✓ Use target weekly vocabulary and grammar forms in speaking contexts. ✓ Demonstrate understanding of the purpose and types of academic presentations. ✓ Demonstrate understanding of the process of developing an academic presentation.

Appendix C: Ethical approval

☆ **ResearchEthicsSystem@glasgow.ac.uk** Inbox - UoG December 19, 2018 at 8:12 AM

Research Ethics Application Approved [A study on EFL students' perception of self-assessment in spea... Details

To: 2310807a@student.gla.ac.uk

Dear Bayan Alghanmi (student),

The following research ethics application has been approved:

Project Title	A study on EFL students' perception of self-assessment in speaking and its effect on achievement, critical thinking, and self-regulation.
Application Number	400180067
Committee	College of Social Sciences
Submitted By	Dr Georgina Wardle

Please log in to the Research Ethics System to download the approval letter from your Application.

This is an automated message. Please do not reply to this email.
If you need additional help, please contact your ethics administrator or visit the IT Services helpdesk.

Appendix D: Participants information sheet (PIS)

University of Glasgow

College of Social Sciences

Participant Information Sheet for Students

Title of Project: **A study on EFL students' perception of self-assessment in speaking and its effect on achievement, critical thinking, and self-regulation.**

Name of Researcher: **Bayan Alghanmi**

Name of supervisors**: Dr. Georgina Wardle** and **Prof. Louise Hayward**

you are being invited to take part in a research study. Before you decide it is important for you to understand why the research is being done and what it will involve. Please take time to read the following information carefully and discuss it with others if you wish. Ask us if there is anything that is not clear or if you would like more information. Take time to decide whether or not you wish to take part.

Thank you for reading this.

You are invited to participate in this research project because the study will be attempting to find out students' perspective of the practice of self-assessment in speaking classes and how they believe its affect their achievement, critical thinking, and self-regulation. The result of this research may be used to provide insight to policy makers on how students perceive the implementation of self-assessment which can help to overcome difficulties and common problems in implementing it with students without experience in self-assessment. It will also provide opportunities for students to reflect on their learning and the areas, critical thinking and self-regulation, that can lead to their progress and achievement which could impact their professional lives in the future.

The participation is voluntary and If you decide to take part in the this study, the study will be conducted for the whole 2nd semester during speaking classes and we will find a mutually convenient time to meet for the interviews at the end of the semester. With your

permission, the speaking activities and interviews will be audio-recorded and I may wish to quote your exact words in my thesis.

Before conducting the study, an introductory PowerPoint presentation will be presented to you with some information regarding the study. The study involves an intervention that will be implemented in speaking classes, which will be held and taught as regular, where you will audio-record your-self during the speaking activity so you can listen to your performance afterward and self-assess your-self with the use of self-assessment sheet after hearing the recording. Also, you will participate and complete two surveys (tests) pre- and post- the intervention at the beginning and the end of the semester, and some of the students will be interviewed at the end of the semester.

All personal information which is collected will be kept strictly confidential. You will only be identified by an ID number, which will be replaced by a code given to you at the beginning of data collection, and you will not be identified in any report or publication.

Confidentiality will be respected unless there are compelling and legitimate reasons for this to be breached. If this was the case, we would inform you of any decisions that might limit your confidentiality.

The results will be included in my thesis and will be completed by 2020/2021. Results may also be available as publication in articles and conference papers. If requested by participants, I will provide them with a summary of the findings. The recordings and data will be stored in my computer and protected with a password or locked in the cabinet in my own accommodation. All electronic or paper copies of data will be retained for 10 years after completion of the project. Please note that the data will not be shared/archived or re-used.

This project has been considered and approved by the College Research Ethics Committee.

For further Information please contact **Bayan Alghanmi**, PhD in Education, School of Education, University of Glasgow, email: **xxxxxxxx@student.gla.ac.uk**
Further information and where to pursue any complaint: this should be the College of Social Sciences Ethics Officer **Dr Muir Houston, email: Muir.Houston@glasgow.ac.uk**

Appendix E: Consent form

University of Glasgow

College of Social Sciences

Consent Form for Students

Title of Project: **A study on EFL students' perception of self-assessment in speaking and its effect on achievement, critical thinking, and self-regulation.**

Name of Researcher: **Bayan Alghanmi**

Name of supervisors: **Dr. Georgina Wardle and Prof. Louise Hayward**

I confirm that I have read and understood the Participant Information Sheet for the above study and have had the opportunity to ask questions.

I understand that my participation is voluntary and that I am free to withdraw at any time, without giving any reason.

I consent to interviews being audio-recorded.

I acknowledge that participants will be referred to by pseudonym.

I acknowledge that follow up interviews might be necessary.

I acknowledge that there will be no effect on my grades arising from my participation or non-participation in this research.

- All names and other material likely to identify individuals will be De-identified (i.e. whereby identifiers are replaced by a code given to participants at the beginning of data collection, to which the researcher retains the key, in a secure location).
- The material will be treated as confidential and kept in secure storage at all times.
- The material will be destroyed once the project is complete.
- The material may be used in future publications, both print and online.

I agree to take part in this research study ☐

I do not agree to take part in this research study ☐

Name of Participant …… Signature ……..

Date ……

Name of Researcher: Bayan Alghanmi Signature ……

Date ……

Appendix F: Self-Regulation of Learning Self-Report Scale (SRL-SRS)

Please read the following statements and choose the answer (on a Likert-scale points) that best describes the way you plan, monitor, reflect, your effort, and Self-efficacy during speaking activities. I would be grateful if you read each statement carefully and provide an answer. Please remember there are no right answers--please describe yourself as you are, not how you want to be or think you ought to be.

	statements	(1) almost never	(2) Sometimes	(3) Often	(4) almost always
Planning	1. I determine how to solve a problem before I begin.				
	2. I think through in my mind the steps of a plan I have to follow.				
	3. I try to understand the goal of a task before I attempt to answer.				
	4. I ask myself questions about what a problem requires me to do to solve it, before I do it.				
	5. I imagine the parts of a problem I still have to complete				
	6. I carefully plan my course of action to solve a problem.				
	7. I figure out my goals and what I need to do to accomplish them.				
	8. I clearly plan my course of action to solve a problem.				
	9. I develop a plan for the solution of a problem.				

	statements	(1) almost never	(2) Sometimes	(3) Often	(4) almost always
Self-monitoring	10. While doing a task, I ask myself questions to stay on track.				
	11. I check how well I am doing when I solve a task.				
	12. I check my work while doing it.				
	13. While doing a task, I ask myself, how well I am doing.				
	14. I know how much of a task I have to complete.				
	15. I correct my errors				
	16. I check my accuracy as I progress through a task.				
	17. I judge the correctness of my work.				

	statements	(1) never	(2) Rarely	(3) Some -times	(4) Often	(5) Always
Evaluation	18. I look back and check if what I did was right.					
	19. I double-check to make sure I did it right.					
	20. I check to see if my calculations are correct.					
	21. I look back to see if I did the correct procedures.					
	22. I check my work all the way through the problem					
	23. I look back at the problem to see if my answer makes sense.					
	24. I stop and rethink a step I have already done.					
	25. I make sure I complete each step					

	statements	(1) Strongly disagree	(2) Disagree	(3) neither	(4) agree	(5) Strongly agree
Reflection	26. I reappraise my experiences so I can learn from them					
	27. I try to think about my strengths and weaknesses.					
	28. I think about my actions to see whether I can improve them					
	29. I think about my past experiences to understand new ideas					
	30. I try to think about how I can do things better next time					

	statements	(1) almost never	(2) Sometime	(3) Often	(4) almost always
Effort	31. I keep working even on difficult tasks				
	32. I put forth my best effort when performing tasks				
	33. I concentrate fully when I do a task.				
	34. I don't give up even if the task is hard.				
	35. I work hard on a task even if it is not important				
	36. I work as hard as possible on all tasks				
	37. I work hard to do well even if I don't like a task.				
	38. If I'm not really good at a task I can compensate for this by working hard.				
	39. I am willing to do extra work on tasks in order to learn more.				
	40. If I persist on a task, I'll eventually succeed.				

	statements	(1) almost never	(2) Sometimes	(3) Often	(4) almost always
Self-efficacy	3. I know how to handle unforeseen situations, because I can well think of strategies to cope with things that are new to me.				
	41. If someone opposes me, I can find means and ways to get what I want.				
	42. I am confident that I could deal efficiently with unexpected events				
	43. If I am in a bind, I can usually think of something to do				
	44. I remain calm when facing difficulties, because I know may ways to cope with difficulties.				
	45. I always manage to solve difficult problems if I try hard enough.				
	46. It is easy for me to concentrate on my goals and to accomplish them				
	47. I can solve most problems if I invest the necessary effort				
	48. When I am confronted with a problem, I usually find several solutions				
	49. No matter what comes my way, I'm usually able to handle it.				

Appendix G: Translation permission

From: **Tynke Toering** tynke.toering@gmail.com
Subject: Re: inquiry regarding (SRL-SRS)
Date: October 24, 2018 at 6:39 PM
To: bayan alghanmi b3.alghanmi@gmail.com

TT

Dear Bayan,

You have my permission to translate the scale to Arabic. Good luck with your research!

All the best,

Tynke

Verstuurd vanaf mijn iPhone

Op 24 okt. 2018 om 18:09 heeft bayan alghanmi <b3.alghanmi@gmail.com> het volgende geschreven:

> Dear Dr Tynke Toering,
>
> Hope this email finds you well
>
> This is Bayan Alghanmi a Phd students. I'm interested in using the "Self-Regulation of Learning Self-Report Scale (SRL-SRS)" in my Phd Thesis. I'll be conducting my research in Saudi Arabia and I would like to ask if I can translate it to Arabic the students first language.
>
> Looking forward to hearing from you
>
> Thank you,
> Bayan

Appendix H: Self-assessment proforma

Speaking Assignment – Week _____ Name_____

Use the self-assessment sheet to plan, monitor, evaluate, and reflect on during speaking activities

CHECK: on a scale from 1 to 5, Think about the Unit Assignment and complete the **Self-Assessment sheet** **5=excellent, 1=not at all**

5=excellent, 4= very good, 3=good, 2=ok, 1= not at all

My plan

_____

Self-assessment						
Criteria			**Work in progress**	**5 point Likert scale**	**Explain your scale (with examples)**	**I plan to improve by**
Criteria		**What do I understand by this criterion**				
Fluency	I spoke without pauses or hesitation.			**5 4 3 2 1**		
Pronunciation	I Show clear pronunciation			**5 4 3 2 1**		
Topic management	The topic was relevant.			**5 4 3 2 1**		
	I covered the topic correctly.			**5 4 3 2 1**		
Content	I used vocabulary from this unit.			**5 4 3 2 1**		
	I used grammar from this unit.			**5 4 3 2 1**		

Grammar (Accuracy)	I used correct sentence structure			**5 4 3 2 1**		
	I used verbs correctly			**5 4 3 2 1**		
	I used nouns correctly.			**5 4 3 2 1**		
	I used adjectives correctly.			**5 4 3 2 1**		
	I used adverbs correctly.			**5 4 3 2 1**		
Vocabulary	I used appropriate vocabulary			**5 4 3 2 1**		
	I used a variety of expressions			**5 4 3 2 1**		
Communicative effectiveness	My Ideas were clear and comprehensive			**5 4 3 2 1**		
	My Ideas were connected and developed.			**5 4 3 2 1**		
Organization & delivery	I managed my time			**5 4 3 2 1**		
	I hold attention of entire class			**5 4 3 2 1**		
	Speaks with fluctuation in volume, clear, audible voice.			**5 4 3 2 1**		

Appendix I: manual analysis of interviews

Interview transcription: S3

Researcher: We are starting the interview about your experience and views of self-assessment and speaking classes. In the beginning we are going to discuss your impression about the experience in general, but before we start I would like to ask you, have you made the audio recordings while speaking?

S: Yes

Researcher: Did you use the self-assessment sheet after listening to the recording or before it or without it at all?

S: yeah I used it after listening to my self.

* she participated in the intervention

Researcher: Okay then, in general, what do you think of self-assessment?

S: yeah it was a nice experience.

* the word nice means good experience and positive attitude.

← positive feeling with self-assessm

Researcher: Can you tell me why do you think it was a nice experience? reasons

S: Well it was like a challenge to me that I had to fully focus on a task and then evaluate myself and find a solution to get better but I really enjoyed it because it was good for me.

* she was fully engaged in her learning which made he concentrate and use cognitive skills (higher) critical thinking skill and metacognitive
* problem solving

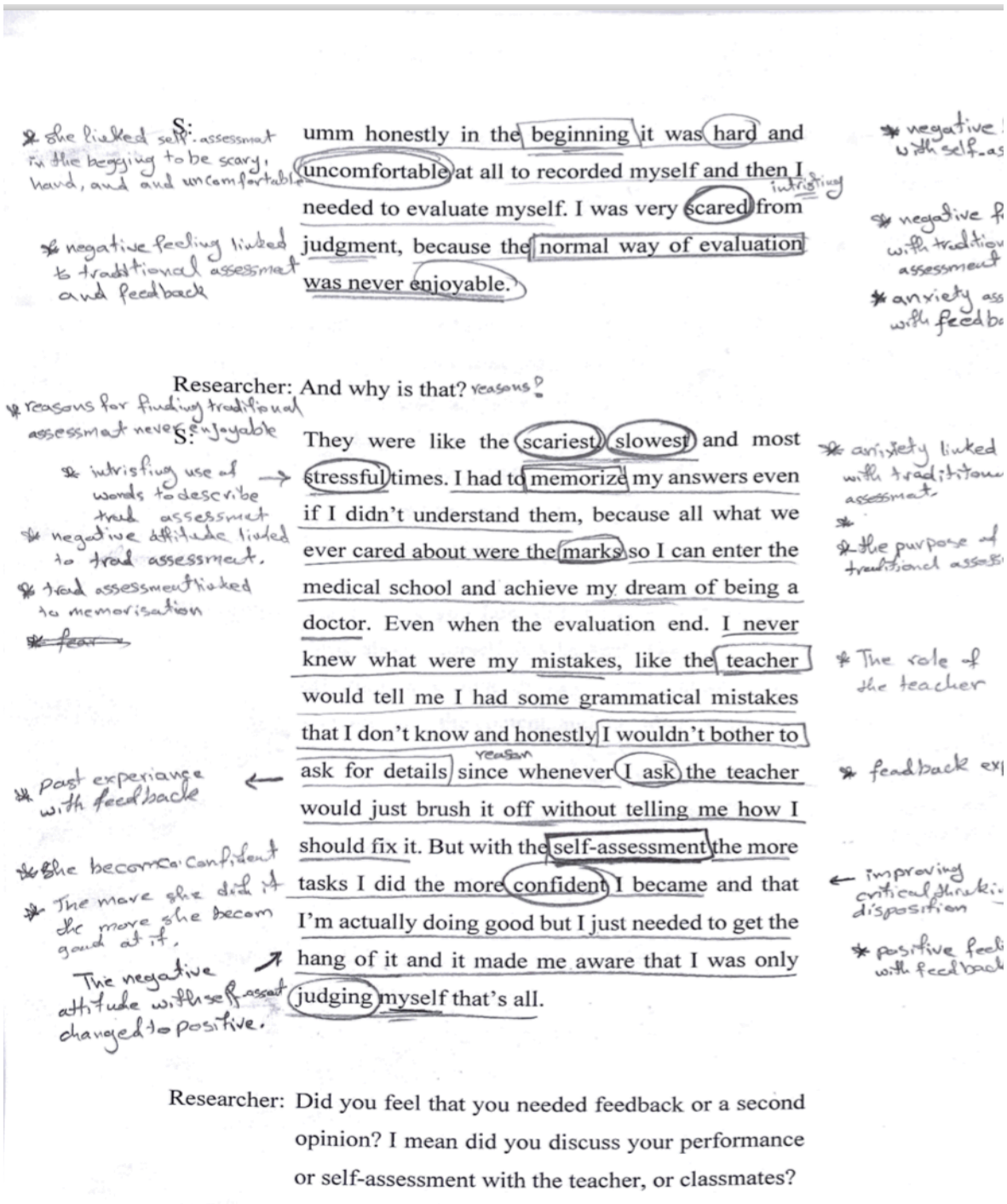
← active in the learning pro

← The use of improving
① cognitive critical thin skills
② metacognitive skills

↖ positive feel with self-ass

↖ The use self assessm

Researcher: how did you feel while using the audio recording and sheet to self assess yourself, were you enthusiastic, confused, what was your general impression?



S: umm honestly in the beginning it was hard and uncomfortable at all to recorded myself and then I needed to evaluate myself. I was very scared from judgment, because the normal way of evaluation was never enjoyable.

Researcher: And why is that?

S: They were like the scariest, slowest and most stressful times. I had to memorize my answers even if I didn't understand them, because all what we ever cared about were the marks so I can enter the medical school and achieve my dream of being a doctor. Even when the evaluation end. I never knew what were my mistakes, like the teacher would tell me I had some grammatical mistakes that I don't know and honestly I wouldn't bother to ask for details since whenever I ask the teacher would just brush it off without telling me how I should fix it. But with the self-assessment the more tasks I did the more confident I became and that I'm actually doing good but I just needed to get the hang of it and it made me aware that I was only judging myself that's all.

Researcher: Did you feel that you needed feedback or a second opinion? I mean did you discuss your performance or self-assessment with the teacher, or classmates?

S. Of course I did, because at first I had no idea if I was on the right track or not, was I even evaluating myself correctly or not but I didn't want to ask anyone in class because one time the teacher threw a marker at me because I made a mistake and when I made a mistake in a group project with my classmates and from that day till the end of the semester they avoided working with me because I ruined it. That's why I prefer asking my close friends for help even though they are not taking the course with me.

Researcher: Let me ask you how did the criteria made you think about yourself as a learner? The assessment criteria here include fluency, pronunciation, topic management, the content, and the other points.

S. well it made me feel like "there's nothing that I can't do or learn if I just work on it and focused on it"

Researcher: Okay, To what extent did the criteria used in the sheet reflect or match the quality that you believe identify a good speaking performance?

S. Everything in it was really important for a good speech.

Researcher: Is there any criterion that you believe should be delete or not included in the sheet?

S: In fact I'm not sure why should we evaluate our time management and topic management. It has nothing to do with our speaking performance and language. Also, at first I honestly thought what's the point to care about adjectives and adverbs in a speech but little did I know they are important. In fact I didn't know the difference between them so I had to search about it therefore I can apply them in the next task.

Researcher: Okay good, is there any criterion that you believe should be added and included in the sheet?

S: umm yes I suggest to add something about body language and face expressions. Because I feel they are important to show a person feelings when talking.

Researcher: So to what extend or how do you think your speaking skill has improved after using self-assessment sheet?

S: oh a lot specially my choice of vocabulary improved I used to only use the words I felt safe with but not anymore and how umm I stopped pausing a lot because I'm feeling more confident to speak right now then before.

Researcher: Were you aware of these aspects before you were engaged in self-assessment?

S: No just to think that I cant focus on one topic and just lost track to what I was saying and had to stop to think what was I talking about again?

The importance and the role of the sheet in raising awareness

* Improving metacognitive knowledge

Researcher: To what extend do you feel you learned something personal gain from self-assessment?

S: I started to plan more in my head before I start talking just to get all my ideas in place and every time I finish I start automatically evaluate my self hmm did I stay on topic? did I used adjectives? was I clear? and then think of ways to be better.

* organisation of ideas.

* She started using SRL strategies.
* planning
* evaluate
* reflecte
* brain storming of ways to improve

* Improving metacognitive strategies
* Improving cognitive skills and critical thinking abilities.

Researcher: To what extend do you think self-assessment helps you discover and plan what and how you would like to practice?

S: for me I think being aware of my own mistakes helps a lot to plan on how to fix them and gets better

* The importance of awareness in improvement
* The purpose and importance of assessment.

* Improving metacog knowledge.
* Improving cogn processes

Researcher: Would you recommend to keep speaking self-assessment for next year?

S: absolutely yes it's a fun challenge to keep on learning with and I like challenges hahah

* It's not an easy task but it's worth the effort.

* positive feeli with self-asses