

## EXPLORING KURDISH EFL STUDENTS' VIEWS ON THE ASSESSMENT PROCESS AT UNIVERSITY LEVEL

Ivan Hasan Murad <sup>1\*</sup>, Sanan Shero Malo <sup>2</sup>

<sup>1</sup> Dept. of English Language, College of Humanities, University of Zakho, Kurdistan Region – Iraq.  
([ivan.murad@uoz.edu.krd](mailto:ivan.murad@uoz.edu.krd))

<sup>2</sup> Dept. of English Language, College of Humanities, University of Zakho, Kurdistan Region – Iraq.  
([sanan.malo@uoz.edu.krd](mailto:sanan.malo@uoz.edu.krd))

Received: 12/ 2024/ Accepted: 03/ 2025/ Published: 03/ 2025 <https://doi.org/10.26436/hjuoz.2025.13.1.1538>

### ABSTRACT:

The current study aims to investigate Kurdish EFL students' views of the assessment process conducted at EFL departments of public universities in the Kurdistan Region of Iraq (KRI). Due to the fact that assessment is the core factor for students' learning, involvement, and evaluation as the only gauge for their progress and development, much attention needs to be given to the assessment process. This study specifically aims at studying the perceptions of the Kurdish EFL students of the criteria including design, administration, purpose, effectiveness and washback, scoring and grading, and feedback of testing and assessment process. Hence, for the purpose of data collection, a questionnaire was administered to 116 students of semesters 3,5, and 7 at the English language departments of some public universities in the KRI during the academic year 2024-2025. Cronbach Alpha was used to analyze the reliability of the items of the questionnaire along with SPSS (version 25) to analyze the mean values of the items and ANOVA was utilized to compare the mean values across the six criteria. Findings indicate significant challenges in the alignment and execution of testing and assessment processes in higher education. While testing and assessment items align with course objectives, they often fail to adequately measure critical thinking and comprehensive language skills. Procedural issues, including unclear instructions, unfair scoring and grading practices, and overemphasis on grading rather than fostering students' progress and engagement, have badly affected the effectiveness of assessments. Additionally, environmental factors such as cheating, unsupportive classroom dynamics, and poor seating quality negatively impact students' performance. A lack of constructive feedback further hinders the development of students' overall skills and learning outcomes. The findings further highlight the need for a holistic approach to assessment that emphasizes student growth, fair evaluation, and the integration of diverse language competencies.

**KEYWORD:** Students' Views, EFL, Test Design and Administration, Scoring, Feedback, Washback.

### 1. INTRODUCTION

#### 1.1 BACKGROUND

Testing and assessment are essential aspects of the educational process as they serve as a crucial tool for determining students' learning and progress, guiding instructional practices, and defining educational outcomes (Brown, 2004). In the context of English as a Foreign Language (EFL), the testing and assessment process measures both language proficiency and communicative competence. This means that in EFL context, assessment takes on additional role by measuring whether or not students have critically and successfully conveyed their intended meaning which is important for their academic success and future opportunities (Putri, Pratolo & Setiani, 2019). Testing is seen as an evaluation tool of retention of knowledge about a particular topic. It is conducted via exams, quizzes, and other standardized tests (Butler, 2021)

that make up the most depended on types of classroom assessment (Frey & Schmitt, 2007). Tests cause negative experience for students as they make anxiety and other negative psychological issues. Assessment in its core educational use reflects a wider and broader concept. It consists of different types including formative and summative, which are both evaluative and supportive for students' development and learning (McMillan, 2011). Therefore, both testing and assessment play crucial but distinct roles in measuring students' achievement and shaping educational practices through which students are better known by teachers. To show the distinction between the two terms: testing and assessment, it is of immense importance to state the mostly used definitions of these two terms.

A very widely used definition of testing is "a method of measuring a person's ability, knowledge, or performance in a given domain" (Brown, 2004: 3). This definition refers

\* Corresponding Author

This is an open access under a CC BY-NC-SA 4.0 license (<https://creativecommons.org/licenses/by-nc-sa/4.0/>)

exclusively to some sort of written tests that measure the insights a person or a student has regarding a particular topic. Typically, this consists of questions, designed activities, or standardized tasks that normally measured in grades or marks. Assessment has been defined by Nitko and Brookhart (2011:4) as "the process of collecting, synthesizing, and interpreting information". This definition denotes that assessment covers a broader concept and takes into consideration a wide range of aspects when it comes to measuring students' potentials, such as observation, peer reviews, presentations, etc. (Brown & Harris, 2016).

In educational discourse, "tests" and "assessment" are interrelated yet distinct concepts. While tests are typically standardized, summative tools aimed at measuring performance, assessment constitutes a broader framework encompassing various evaluative methodologies. It includes both formative and summative approaches designed to monitor learning progress, competency development, and skill acquisition (Brown & Abeywickrama, 2019). This distinction is essential for maintaining conceptual clarity and inclusivity in educational practices, as many evaluative methods, despite not being classified strictly as "tests," still serve critical assessment purposes. Carless (2015) highlights the broader role of assessment beyond testing, emphasizing its importance in delivering feedback and fostering learning improvement. Therefore, in this article the term 'test and assessment' has been used to give more clarity and inclusivity to the testing and assessment practices followed in the current Kurdish EFL contexts.

In the higher education (HE) context of KRI, particularly in EFL context, testing and assessment complete each other in measuring students' progress. Black and Wiliam (2009) state that testing and assessment involve ongoing evaluations in which feedback is provided to students throughout the learning process. This is referred to as formative assessment which is rarely graded in our HE context. Most EFL departments in HE universities in KRI teach both literary and language learning and linguistic subjects. In most departments, the first two years heavily focus on improving learners' language skills of the learners. The third (juniors) and fourth (seniors) year students are taught both literary and linguistic subjects which require a range of different assessments that enable teachers to fairly evaluate their students' knowledge and performances. Black and Wiliam (1998) stated that the assessment that is conducted in the classroom is paralyzed as it "encourages superficial and rote learning" (p.17) and that focus is more on the graded items which students memorize and soon forget. This is argued to be the typical situation of the EFL context in Kurdistan Region HE EFL departments wherein graded tests (mid-term tests, quizzes, and final or end of semester tests) are given priority over performance assessment and demonstration of learning. Such a predicament would rather lead the students to construct a negative perception of the whole assessment process as it overlooks their abilities and disregards their performance in the classroom. Moreover, students' perception of the assessment process and the learning environment have a profound impact on their academic progress and educational practices (Brookhart, 2017). It

has been noted that the learners who have positive perceptions of their learning and assessment process are more likely to boost their learning and achieve higher academic outcomes (Paechter et al., 2010). This further supports Brookhart's (2004) argument that the way students' perceptions of the assessment process can impact their enthusiasm and attitude towards learning. Therefore, it is crucial to investigate students' attitudes and perceptions of the assessment process as it plays a very critical role in shaping the way they view learning, engagement, motivation, the value of education, and their performance (Carless, 2015).

## 1.2. STATEMENT OF THE PROBLEM

The problem highlighted in this study lies in the fact that the assessment process conducted at EFL departments does not meet the expectations of the students. This might lead students to become unengaged with the process of learning and unresponsive to the educational environment, particularly due to the types of assessments that are heavily depended on. This is because the focus could be limited to the summative types of assessments such as standardized final exams and mid-term tests, which only account for evaluating how much learning is achieved at the end of a term, course, or unit of study (Hattie & Clarke, 2019). Moreover, such an assessment process might lead students to view assessment as only high-stakes causing high stress and fostering anxiety rather than utilizing the best assessment practices through which students' overall potentials can be considered. In addition, the heavy dependence on summative assessments might result in biasness in grading-whether or not due to subjective grading or inconsistency in dealing with the students, a procedure that might further complicate the students' perceptions of fairness throughout the whole process (Hughes, 2003).

Another essential concern lies in the design and administration of assessments tools. Both the designing of tests and their administration play a pivotal role in shaping students' perceptions of the assessment (Cheng, 2017). Poorly designed assessments which might lack face validity and do not properly align with the course objectives or real-life language use can have undue implications on the outcomes of the assessments conducted (Hughes, 2003). According to the best knowledge of the researchers, no studies have been conducted to investigate Kurdish EFL students' attitudes of the assessment process in terms of test and assessment design, administration of assessments, purpose of assessment, effectiveness of assessment, scoring and grading practices, feedback, and washback of the assessment process. On this basis, it is of dire need to study the perceptions of the Kurdish EFL students of the assessment process conducted at EFL departments at public universities in the KRI.

## 1.3. SIGNIFICANCE OF THE STUDY

The significance of the current study lies in its ability to contribute to the broader field of assessment in HE particularly in KRI, especially within the context of EFL education. Student attitudes towards the assessment process play a pivotal role in shaping their academic engagement, motivation, and overall learning outcomes,

highlighting the importance of examining how they interpret and respond to various assessment methods. In Kurdish HE EFL contexts, assessment practices should align with students' unique needs and expectations to foster meaningful learning experiences. Beyond merely evaluating proficiency, effective assessment serves as a tool for formative feedback, guiding students in their linguistic development (Carless, 2015). Furthermore, cultural and institutional factors significantly influence students' attitudes toward testing and grading, necessitating the implementation of assessment designs that are contextually appropriate (Brown & Harris, 2016). Challenges such as large class sizes, limited technological resources, and policy constraints further emphasize the need to explore student perspectives in order to develop assessment strategies that balance validity, reliability, and practical feasibility.

#### 1.4. RATIONALE FOR THE CONDUCTION OF THE STUDY

The rationale behind conducting the present article is that despite the fact that a number of studies have been conducted to investigate the views of the Kurdish EFL students (Mahmood & Ghaleb, 2024; Qadir et al., 2023), these studies have not tackled the above mentioned variables in the assessment process, namely test and assessment design, administration, purpose, effectiveness, scoring and grading practices, feedback, and washback effect of the assessment process. Moreover, the results of the study might have practical implications for teachers, administrators, and decision makers in the context of HE. Consequently, it strives to bridge a gap in the existing literature by providing unique insights into the procedures adopted to improve assessment practices and students' outcomes.

#### 1.5. AIMS OF THE STUDY

##### **This study aims to**

1. investigate Kurdish EFL students' views of the design, administration, and purpose of the testing and assessment process at university level.
2. explore Kurdish EFL students' opinions regarding the effectiveness and washback of the assessment process, the scoring and grading practices, and the feedback process in testing and assessment.
3. pinpoint the statistical difference between the six criteria (design, administration, purpose, effectiveness and washback, scoring and grading, and feedback) in terms of students' views.

#### 1.6. RESEARCH QUESTIONS

Based on the above statement of the problem, the significance of the study, and the rationale of the study, the following research questions have been forwarded:

1. What are Kurdish EFL students' perceptions of the design and administration of tests and assessments at HE institutions?
2. What are the students' perspectives of the purpose and effectiveness of tests and assessments at their HE institutions?
3. How do students evaluate the scoring and feedback process provided on their assessments?

4. What do students think about the overall perceived washback effect of the assessment process?

5. Is there a statistically significant difference between the investigated criteria (design, administration, purpose, effectiveness and washback, scoring and grading, and feedback) of the assessment process according to students' views?

## 2. LITERATURE REVIEW

Assessment is a broad term which includes testing and a plethora of assessment methods. It consists of formative assessments that take place during the period of learning and provide feedback immediately that helps students improve and build on their knowledge alongside the summative assessment methods which provide final evaluation a learner's achievement at the end of an instructional period (Dixson & Worrell, 2016). Many scholars argue that assessment is not merely one way of evaluation but a systematic process that incorporates a number of techniques, methods, and strategies that provide a comprehensive evaluation of students' learning and abilities (Brookhart & Nitko, 2019; McMillan, 2011; Popham, 2017; Stiggins & Chappuis, 2017; Wiliam, 2018). Testing, on the other hand, is a proper technique of evaluation that is utilized to assess students' proficiency or knowledge (Fulcher, 2019). Testing is often associated with increased students' anxiety and stress, particularly when it is done in high-stake environments where outcomes influence students' progression and knowledge (Jerrim, 2022). It is argued that testing is primarily used for summative evaluations and this is the case in Kurdistan Region universities where testing is more valued than other forms of evaluation techniques (Fulcher, 2013). Consequently, students' negative perceptions are formed due to the heavy use of tests primarily as summative evaluation techniques.

Perception is defined as the way in which individuals interpret and realize the world around them by basing their interpretation on their past experience, expectations, and cognitive frameworks (Dörnyei, 2007). At EFL contexts, students' perceptions are highly valued and crucial in defining the way they approach learning and the way they engage in the assessment process. For McMillan (2011), students who see assessments as fair, transparent, and effectively aligned with the learning objectives are more engaged and have higher chances of achieving positive learning outcomes. It is essential to indicate that in this study, perception and view have been used interchangeably.

### 2.1. TYPES OF ASSESSMENT

In the EFL context, especially the HE context at universities in the KRI, the assessment process is referred to the practices and assessment techniques that teachers implement in the classroom. These practices are mainly designed to measure the progress students achieve so that teachers can depend on to shape their pedagogical choice. Recent literature has underlined several types of classroom assessment, including formative assessment, summative assessment, and alternative assessment (Mngomezulu et al., 2022), which have distinct educational purposes (Brown & Abeywickrama, 2019). According to Black and

Wiliam (2018) assessments such as quizzes, assignments, in-class activities, peer assessments, interactive questions, all serve as a gauge for identifying students' progress, strengths, weaknesses, and provide feedback for improvement in the learning process. This aligns with the study of Mngomezulu et al. (2022) who state that quizzes, assignments, in-class activities, peer assessments, and interactive questions are essential for students' engagement in the classroom, because they promote self-regulatory learning behaviours through which students can monitor and adjust their potentials (Black & Wiliam, 2009).

Summative assessments, on the other hand, are those types of assessments that are usually conducted at the end of a term, course, or unit of study (Basera, 2019). Examples of these assessment include term papers, national exams, final exams, mid-term tests, and final year projects that provide a comprehensive idea about students' achievement (Black, 1993). This type of assessment plays a significant role in the global education system especially in Iraq and KRI (Qadir et al., 2023), because it is regarded as the primary decision for the certification, accountability, and provision of evidence for students' progression (Broadfoot, 2007; Smith & Fey, 2000). It could be argued that this type of assessment is the most influential one in engaging students in the learning process because of the significant role it plays in deciding students' future. However, according to Boud and Falchikov (2006), due to the fact that it gathers insights about students learning and achievement in a single snapshot of demonstration of knowledge, scholars have scrutinized it for not providing a holistic picture of the potentials students are endowed with. Moreover, summative assessment promote memorization rather than helping students learn in a deeper level.

## **2.2. ASSESSMENT PROCESS**

### **2.2.1. TEST AND ASSESSMENT DESIGN**

In EFL context, one of the most essential aspects of assessment is the test design as it entails a number of crucial factors that any test should fulfil. Construct validity is one of the factors that ensures that the test and its items measure what it intends to measure (Fulcher, 2010). In the past, validity of an instrument was determined by what it was used to assess (Lado, 1961; Brown & Abeywickrama, 2019). This perspective on the validity of a test was soon replaced as rigorous validation processes were introduced (Al-Wadi, 2020) and necessitated the alignment of test items with the content of the material and the theoretical constructs (Pond, 2019). According to Murphy et al. (2023), the breadth of the material and curriculum must be reflected in the test items, thus enabling the test to measure learners' abilities. On this basis, the testing instrument would be reliable and authentic yielding clear results of the potentials of the learners' use of language in non-testing situations (Bachman & Palmer, 1996). The factor regarding the breadth of the material also aligns with Hultgren et al.'s (2022) study in terms of test formatting and instructional difficulty. Hultgren et al. (2022) state that tests are ought to be sufficiently challenging for the learners to effectively measure their abilities and not to cause any undue psychological problems, but rather they should enable critical thinking skills which foster learners'

evaluation, analysis, and higher-order thinking abilities. Bennett (2011) adds to this stating that for test to gather authentic insights about the students' progress, it should follow the evidence-centered design (ECB) framework. This framework works toward shifting the purposes of tests and exams to become more evidence-gathering processes: i.e., the test items should reflect the content they are trying to measure.

### **2.2.2. TEST AND ASSESSMENT ADMINISTRATION**

According to Harlen (2021), test administration consists of a number of aspects such as, planning, organization, and administration of tests to students, clear communication of instruction and assessment expectation such as test format and deadlines to decrease students' anxiety level. Also, the environment of the test settings which involves the physical arrangement of the facilities, such as the heating and cooling levels, the provision of suitable seating in terms of quality (Cheryan et al., 2014), and for those individuals with special needs is all conducive to students' successful performance (Fitriyah et al., 2022). At many universities, class size and limited facilities can affect the proper administration of tests. Furthermore, effective authoritarian control characterized by effective invigilation during tests results in the provision of a setting equitable to all students in terms of security and control (Bachman & Damböck, 2018; Crossley, 2022; Van Bergen & Lane, 2014). This is further confirmed by Hughes (2003) who states that invigilator's conduct might cause discomfort to the test taker, but successful management and proctoring result in positive student performance ensuring a fair and transparent administration of the test. Likewise, excessive noise level can negatively affect students' performance. According to Klatte et al. (2013), it is essential that tests are conducted in a quiet environment so that students' concentration is not disrupted.

### **2.2.3. PURPOSE AND EFFECTIVENESS OF TESTS AND ASSESSMENT**

In educational settings, testing and assessment serve a number of purposes including administrative, pedagogical, motivational, and institutional. Depending on the type of assessment, whether formative or summative, the purpose for conducting it is multifaceted. According to Carless (2015), tests and assessments that are formative in nature provide constructive learning feedback and guide the instructional and learning process. Fulcher (2019) agrees with Carless and states that in addition to guiding instruction, formative assessment also provides students with their progress (Tsagari & Vogt, 2017), and shapes curriculum and pedagogical practices, thus leading to a refinement in the students' learning strategies (Black & Wiliam, 2018). This makes assessment effective as it drives students to employ self-regulated learning process and dynamic assessment to improve their comprehension and critical thinking skills (Brown, 2004). Such a concept and purpose is also referred to as 'assessment for learning' (Manitoba Education, Citizenship & Youth, 2006), for it works towards finding out effective tools for measuring what students know and can do and how they can further improve their knowledge.

Tests and assessments that are utilized for the purpose of grading and decision-making regarding students'

achievement (Taras, 2005) are referred to as summative assessment. Such assessment is conducted at the end of a term, course, or academic year. These assessments normally refer to the measures taken to indicate whether or not students have met the requirement of a curriculum or confirm what students have learned (Manitoba Education Citizenship and Youth, 2006; Tsagari & Vogt, 2017). They are used as a benchmark for determining the achieved level of students for the purpose of decision-making regarding students' achievement and as a proof of course completion. It is argued that such assessments, if well-structured and designed, can encourage students' critical involvement (Brookhart, 2013), leading to effective performance and active learning. Conversely, tests and assessments that are not well aligned with course objectives and designed with poorly written items might cause psychological problems for students and lead to surface learning. According to Biggs and Tang (2011), such assessments drive students to focus on what they expect to come in the tests, hence focusing on memorizing bits of information rather than being involved critically in the learning and comprehension of information.

#### 2.2.4. SCORING AND GRADING PROCESS

One of the most crucial and critical aspects in testing and assessment in HE especially Kurdistan HE institutions is scoring and grading. Scoring constitutes the systematic assignment of numerical values to student responses, predicated on established evaluative criteria. This process yields an objective quantification of performance across discrete assessment components, including individual test items, assignments, or comprehensive assessments (Brookhart, 2013). Primarily focused on the evaluation of specific student outputs, such as responses to selected-response items, constructed-response essays, or problem-solving tasks, scoring precedes the aggregation of these numerical values into broader measures of achievement.

Grading, conversely, represents a more comprehensive evaluative procedure, encompassing the synthesis and interpretation of accumulated student scores into a summative judgment (Brookhart, 2013). Typically expressed as letter grades or percentage scores, grading integrates data from multiple assessments, assignments, and participation metrics to provide a holistic representation of student academic performance (Guskey & Brookhart, 2019). In contrast to the strictly objective nature of scoring, grading may incorporate elements of subjective professional judgment, including considerations of student effort, demonstrated improvement, and the instructor's informed evaluation.

As earlier mentioned in the section on research problem, due to the fact that the value of the majority of the grades in EFL departments in KRI is given to the summative assessment for decision-making regarding the future of the students, special attention should be given to the process of scoring and grading. Many studies argue that fairness in scoring and grading influences the way students perceive the assessment process as a whole (Brown & Harris, 2016). Brookhart (2017) states that demonstrating fairness in the process of scoring and grading motivates students and engages them in the educational process, as it reflects integrity in the assessment process, minimizes biasness, and provides a feeling of relaxation for students. A serious

issue in scoring and grading is subjectivity. It is argued that students at EFL departments are very worried about this phenomenon especially in testing their productive skills (writing and speaking) as they see that the judgement in grading these skills can vary from a teacher to another (Fulcher, 2019). In the same manner, Jonsson and Svingby (2007) advocate the use of rubrics in evaluating students' work as they provide fair evaluation, minimize bias and make sure that consistent evaluation is practiced by different teachers. This, also, ensures that transparency in the grading practices is achieved and students are aware of how their answers are graded (Moss et al., 2006; Sadler, 2009). Sadler (2009) further states that the scoring and grading process must take into consideration students' needs and diverse learning styles and abilities. Such professional practices in scoring and grading promotes equity amongst students, ensures their engagement in the learning process, and gives them all a sense of inclusion and an equal opportunity for success (Tierney, 2014).

#### 2.2.5. THE FEEDBACK PROCESS

In any EFL context, the purpose of feedback takes several forms and is provided for a number of reasons, both constructive and destructive. Effective feedback which is provided in a timely manner can enhance students' learning (Voinea, 2018), as it gives them the opportunity to make immediate necessary correction that have paramount positive implication for their overall learning (Hattie & Timperley, 2007). Positive and supportive feedback provides students with suggestions that can boost their improvement. According to Black and Wiliam (1998), providing constructive feedback and avoiding negative or too judgemental feedback helps improving their performance. Such feedback, further, encourages a growth mentality amongst students, hence resulting in a constructive progress that bridges the gap between current performance and the improvement results in the learning process (Brookhart, 2008). Therefore, the constructive feedback, known as formative type of feedback, is proven to facilitate learning and enhance students' engagement and motivation to learn (Aslam & Khan, 2020). On the other hand, destructive feedback, or the feedback that is particularly provided to pinpoint negative aspects in students' work, can have devastating effect on the students' learning, focus, and overall progress.

Many scholars purport that feedback that is provided to students must be fair, clear, and easy to understand and follow (Black & Wiliam, 1998; Sadler, 2009). Due to the fact that feedback can lead to anxiety, hence causing students to become disengaged in the learning process, eventually hindering learning and understanding, it should be clear and easy to understand (Dabiri, 2018). Therefore, it is of immense importance that students are provided with clear feedback, a comprehensive explanation of the scoring process, and the rationale for the feedback provided for the students to understand their strengths and the areas that they need to further work on (Obilor, 2019). To further support this argument, Lin et al. (2023) encourage every teacher to utilize rubrics in providing feedback to their students for their clear evaluation and criteria which support students with clear expectations of what their teachers want as well as providing clear outlines for the areas that need improvement.

### 2.2.6. WASHBACK EFFECT OF TESTS AND ASSESSMENTS

According to Messick (1996), washback or backwash assessment is the impact, whether positive or negative, an assessment method may have on the behaviour of the teachers or the students. Many scholars have highlighted the positive influence of washback by stating that it occurs out of an effective utilization of assessments which in turn enhances effective learning and boosts students' critical involvement (Alderson & Wall, 1993; Cheng & Green, 2007). Cheng and Green (2007) further maintain that if assessment tools are effectively aligned with the learning outcomes and objectives of the course, they can drive students to focus more on their curriculum, become motivated to learn, and eventually improve their language skills. However, in contexts such as Kurdish EFL departments, where there are class of big sizes (Murad, 2015), it could be challenging for teacher to provide students with individualized feedback and employ effective assessment methods. This could lead them to heavily depend on standardized exams, especially those that promote memorization of grammar, vocabulary, and language chunks. Hence, it is argued by many that teachers, in contexts where there is a heavy dependence on high-stakes testing, prioritize test preparation leading to a focus on a narrow curriculum and ignoring other aspects such as communicative and interactive skills (Hughes, 2003; Cheng, 2005).

Morrow (1986) states that the relationship between assessment and curriculum is based on washback validity. Messick (1996) further reports the concept of washback as an indicator of the consequential dimension of construct validity of assessments, linking positive washback with authentic assessment. In the same vein, Weir (1990) states that a test based on communicative aspects of language and designed to address such skills, promotes positive washback as it is closely linked to authentic language learning and use.

On this basis, to mitigate the negative washback of assessments, scholars (e.g., Weir, 2005) have suggested well-adjusted testing and assessment methods through the use of a variety of different assessment approaches including formative and summative to evaluate students in a fair and authentic manner. Furthermore, Cheng and Green (2007) endorse the effective alignment of assessments with course objectives and goals to provide effective judgement and guide teaching and learning.

### 2.3. PREVIOUS STUDIES

Many studies have been conducted to investigate students' perceptions on some particular assessment and evaluation types and techniques at HE level. However, studies on the perceptions of students on testing and assessment process from the perspectives of design, administration, purpose, effectiveness, scoring and grading, feedback, and washback process have not been widely conducted. However, several studies have been carried out. They serve a similar purpose are summarized as follows:

A very recent study on "Kurdish EFL students' perceptions of summative and formative assessment at Salahaddin University" has been conducted by Mahmood and Galeb (2024). The main aims of the study were to (1) examine

the perceptions of Kurdish EFL students of summative and formative assessment, (2) check if improvement is needed in the assessment system to better meet students' needs, and (3) determine which assessment type has more positive impact on students' learning. For this purpose, a mixed method was utilized to collect data from 542 participants who responded to a questionnaire and 30 interviewees who answered some qualitative questions for the study. Results revealed that Kurdish EFL students at Salahaddin University are not satisfied with the assessment system followed and that they believe it needs to be changed by incorporating more formative assessment strategies.

Another recent study by Wang et al. (2023) was conducted to investigate the perceptions of undergraduate and graduate students in Taiwan and the USA of formative and summative assessments. One of the main purposes was to study the differences between graduate and undergraduate students' perceptions of assessment, and analyze the cross-cultural differences in perceptions of the assessments used. In this study, a questionnaire based on the theoretical framework and previous studies was administered to 349 undergraduates and graduate Taiwanese students and 97 American undergraduate and graduate students. Findings in this study revealed that, according to the students' perceptions, Taiwanese teachers were relying on attendance, classroom participation, homework, and quizzes/exams in assessing their students whereas, the U. S. assessment practices focused more on learning diaries, essays, presentations, and projects. Taiwanese students had positive perception of self-assessment practices, peer-assessment, and varied methods of assessment. On the other hand, U.S. students perceived assessments as tools for improving the quality of teaching and assessment strategies.

A study entitled "Student Perception towards Mandated Assessments," was conducted by Woolever in 2019. The study attempted to gain information concerning students' views on the mandated tests to determine their perceived value in the areas of (1) improvement, (2) external attribution, (3) affective benefits, and (4) irrelevance. The study was conducted with 360 ninth and tenth grade from five high schools in the USA. The results showed that 9<sup>th</sup> and 10<sup>th</sup> graders disagreed with the value mandated tests have on their improvement, external attribution, and affective benefits. Moreover, the mandated tests in their views were irrelevant. Females participants viewed assessments as unfair and not a good measure for the quality of the school and learning. Moreover, English language students perceived the mandated test as irrelevant when compared to other language learning students.

Another study conducted to investigate "learner's perceptions of assessment and testing in EFL classrooms in Albania," by Vavla and Gokaj (2013). The study aimed at checking EFL students' perceptions of assessment and testing. A mixed method approach using a questionnaire and interviews was utilized to collect data from Albanian EFL learners. The results indicated that Albanian learners had no decision on the process of assessment and testing and that it was a purely teacher's role in the process of



education. It is also revealed that assessment and testing are demotivating in terms of learning tools.

### 3. METHODOLOGY

#### 3.1. RESEARCH DESIGN

A descriptive survey design is utilized in this study to investigate the perceptions of Kurdish EFL students towards the assessment process used at EFL department of the public universities in the KRI. Specifically speaking, this study examined the students' perception of the testing and assessment design, administration, purpose, effectiveness, scoring, feedback, and washback of the assessment process. According to Babbie (2021), descriptive surveys are effective in collecting large amounts of data from a wide population area. Moreover, it is used to measure beliefs, perceptions, and attitudes (Sekaran & Bougie, 2020), as well as a characteristic of a particular group of people which helps in decision-making, identify trends, and builds plans for future (Cooper & Schindler, 2014).

#### 3.2. PARTICIPANTS AND SAMPLING

In the current study, 116 undergraduate Kurdish EFL students took part. They were from a diverse number of public universities in the KRI representing a diverse group in terms of age (18-24 years old), study stage at university, and gender. According to Mertens (2019), such a diverse group provides valuable insights about the phenomena under investigation for the varied experience they have concerning the testing and assessment process. Since the students were already set into their classes and stages, a purposive sampling method was employed as it provides the opportunity to select a group of people or participants based on their relevance to the research area (Cohen et al., 2017). Purposive sampling is a type of non-probability sampling which refers to the selection of participants according to specific criteria that they have and are relevant to the research questions (Dörnyei, 2007).

#### 3.3. DATA COLLECTION (PROCEDURES AND MATERIALS)

This study employed a Likert scale questionnaire, which is a popular instrument used by many scholars in the educational field for collecting data on attitudes and perceptions in educational settings (Joshi et al., 2015). According to Sauro and Lewis (2016), Likert scale questionnaires are best tools for collecting data from a wide population on the condition that the researcher ensures its validity and reliability. The Likert scale questionnaire utilized in this study was especially constructed for the purpose of data collection concerning the Kurdish EFL students' views of the testing and assessment process in terms of design, administration, purpose, effectiveness and washback, feedback process, and scoring and grading process. The majority of the items of the questionnaire were adapted from the ideas of Bachman and Palmer (1996 & 2010), Fulcher and Davidson (2007), and Brown and Abeywickrama (2019) studies.

After the questionnaire was constructed, it was given to a panel of jury of 10 experts from a variety of universities in KRI, specialized in Applied Linguistics. Then, it was

evaluated by the jury for validity (face, content, and construct validity) and reliability.

Afterwards, the questionnaire was edited based on the modifications suggested by the jury members. It was then converted into an online survey via Google Form. Then, the link for the survey was shared with 15 EFL students from different stages at the university of Zakho/College of Humanities/Department of English Language for piloting purposes. Based on the feedback given by the piloting sample, a number of items were edited as they were confusing. Then, after 15 days, another round of piloting was performed to ensure that the revisions were effective. Brace (2008) states that by piloting an instrument, the validity and reliability of the tool is increased and it becomes more practical.

Following this, the survey was sent to a number of universities in the KRI, namely University of Duhok, University of Salahaldin, Soran University, Garmian University, University of Raparin, University of Halabja, University of Koya, University of Sulaimani, and University of Zakho for a period of two weeks after which a representative number of 116 responses was received.

#### 3.4. ETHICAL CONSIDERATIONS

The participants were fully informed about the purpose of conducting the study and were told that the participation is totally voluntary following the ethical practices of the research conduction in education (Mertens, 2019). This was to make sure that students are aware of their right as to either participate or not and that they could withdraw at any time they wished. Moreover, they were informed that their data would be kept confidential and stored securely to prevent unauthorized access (Babbie, 2021).

### 4. RESULTS

The survey utilized in the study consisted of six criteria to investigate the testing and assessment process, namely Design (D), Administration (AD), Purpose (P), Effectiveness and Washback (EW), Scoring and Grading (SG), and Feedback (F). Each criterion consisted of six items. Table 1 displays the results of the reliability analysis of the items using Cronbach's alpha.

**Table (1)** Reliability check Using (Cronbach's  $\alpha$ )

Reliability Statistics			
Cronbach's Alpha		N of Items	
.906		6	

  

Case Processing Summary			
		N	%
Cases	Valid	116	100.0
	Excluded <sup>a</sup>	0	.0
	Total	116	100.0

a. List wise deletion based on all variables in the procedure.

As it is known, Cronbach's alpha evaluates a scale's internal consistency or dependability of a questionnaire or a survey. It evaluates if a group of items or categories measures the same underlying concept.

Therefore, as shown in Table 1 above, an alpha value of 0.906 indicates that all the items of the six criteria used in the survey are highly consistent and well-correlated and that each criterion effectively measures a single underlying construct.

Table (2) shows the mean value for the items in the Test and Assessment Design.

**Table (2)** Test and Assessment Design (D)

No.	item	N	Mean	Std. Deviation	t	p-value
1	The assessment tasks are clearly aligned with course objectives.	116	3.24	1.092	2.380	0.02
2	The instructions are easy to understand.	116	2.91	1.179	-0.866	0.39
3	The difficulty level is appropriate.	116	3.09	1.154	0.805	0.42
4	The test help demonstrate learning.	116	2.86	1.062	-1.398	0.16
5	Time for test completion is sufficient.	116	3.14	1.243	1.195	0.23
6	Tests and Assessment are clear and related to content.	116	3.27	1.122	2.566	0.01
7	The design encourages critical thinking skills.	116	2.96	1.122	-0.414	0.68
8	Tests and Assessment cover all course material.	116	3.11	1.045	1.155	0.25

As indicated in Table 2, statistically significant differences are found in Item 1 with mean value 3.24 and  $p=0.02$ : *the assessment tasks are clearly aligned with course objectives*, and Item 6 with mean value 3.27 and  $p=0.01$ : *tests and assessments are clear and related to course content*. Therefore, it was found out that Kurdish EFL students' perception on the alignment of their tests and assessments with course objectives and the clarity and relevance of the testing items and assessment is positive. However, their perceptions on the remaining items was not significantly strong as the p-value was higher than the level of significance 0.05. This suggests that the neutral perception regarding the rest of the Items, 2, 3, 4, 5, 7, and 8, indicate that intervention is required in these areas for the purpose of improving the quality of test and assessment design within the Kurdish EFL context.

**Table (3)** Test and Assessment Administration (AD)

No.	item	N	Mean	Std. Deviation	t	p-value
1	Invigilation rules negatively affect performance.	116	2.58	1.073	-4.242	0.00
2	The desk I sit on is comfortable and causes no distraction.	116	2.68	1.374	-2.500	0.01
3	Administrative procedures are clear and facilitative.	116	3.13	1.146	1.215	0.23
4	Students are informed about test and assessment schedule.	116	3.61	1.170	5.632	0.00
5	Testing and assessment environment is supportive.	116	3.09	1.370	0.745	0.46
6	Effective conduct of invigilators during tests and assessments.	116	3.34	1.215	2.979	0.00
7	No chance for cheating during tests and assessments.	116	3.03	1.367	0.204	0.84
8	Testing and assessment rules are given to students in advance.	116	3.63	1.115	6.077	0.00

Table 3 displays participants' views of the Test and Assessment Administration (AD). Item 1 was intentionally negatively worded to make sure whether or not participants carefully read and respond to the items. As it can be noted, Kurdish EFL students' perception of Item 1 was negative with mean value 2.58 and  $p=0.00$ , indicating that they disagree on the item: *invigilation rules negatively affect performance*. However, for Item 2, with its mean value 2.68 and  $p=0.01$ , the participants suggested concerns about whether or not the desk they sit on is comfortable and causes no distraction. This indicates that their perceptions regarding the quality of the desk they sit on is generally negative and that it causes distraction during the tests and assessment process. On the other hand, an agreement was indicated for Items 4 and 6: *students are informed about test and assessment schedule* and *testing and assessment rules are given to students in advance* with mean values 3.68 and 3.63 respectively. Table 1 also indicates neutral responses for Items 3, 5, 6, and 7. This means that Kurdish EFL students neither agree nor disagree to the items: *testing and assessment administrative procedures are clear and facilitative*, *testing and assessment environment is supportive*, *effective conduct of invigilators during tests and assessments*, and *no chance for cheating during tests and assessments*.

Overall, it could be deduced from Table 3 that key strengths lie in the clear communication of schedules, effective conduct of invigilators, and advance provision of testing and assessment rules. Whereas, the potential area of improvement is 'providing comfortable desks to sit on to enhance test and assessment performance' and needs to



be taken into consideration when administering tests or assessments.

**Table (4)** Purpose of Tests and Assessment (P)

No.	item	N	Mean	Std. Deviation	t	p-value
1	Tests and assessment help in measuring students language ability.	116	3.04	1.219	0.381	0.70
2	Tests and assessment measure students' progress.	116	3.00	1.030	0.000	1.00
3	Tests and assessment check students' achievement.	116	3.28	1.094	2.801	0.01
4	Tests and assessment identify students' weak and strong points.	116	3.27	1.122	2.566	0.01
5	Tests and assessment are only for grading purposes.	116	3.18	1.213	1.607	0.11
6	Tests and assessment helps teachers enhance teaching.	116	3.13	1.191	1.170	0.24
7	Tests and assessment monitor students' engagement.	116	3.17	1.041	1.784	0.08
8	Tests and assessment provide supportive feedback.	116	2.96	1.122	-0.414	0.68

It can be noted in Table 4, which presents students' views towards the purpose of tests and assessments at EFL departments, that statistically significant results could be found for Item 3: *tests and assessment check students' achievement* with mean value 3.28 and  $p=0.01$  and Item 4: *tests and assessment identify students' weak and strong points* with mean value 3.27 and  $p=0.01$ . This indicates that the participants agree that tests and assessments effectively evaluate their achievement and help identify their strong and weak points.

Neutral views can be observed in the rest of the items 1,2,5,6, and 7, as their mean values are 3.04,3.00,3.18,3.13,3.17 and 2.96 respectively and their p-value exceeds 0.05. These findings indicate that areas such as items 1,2,5,6, and 7, stated in Table 4 point to potential areas for improving the efficacy of testing and assessment process in educational settings in the Kurdish EFL context.

**Table (5)** Test and Assessment Effectiveness and Washback (EW)

No.	item	N	Mean	Std. Deviation	t	p-value
1	Test and assessment result in better learning.	116	3.09	1.123	0.826	0.41
2	Test and assessment cover range of language skills.	116	3.10	1.042	1.070	0.29
3	Test and assessment reflect course content.	116	3.29	1.004	3.143	0.00
4	Test and assessment help identify areas to focus on.	116	3.18	1.035	1.884	0.06
5	Test and assessment influences the way students study.	116	3.15	1.136	1.389	0.17
6	Test and assessment motivate students to study effectively.	116	3.14	1.208	1.230	0.22
7	Test and assessment have positive impact on my learning strategies.	116	3.03	1.198	0.233	0.82
8	Test and assessment engage students in class activities.	116	3.12	1.195	1.087	0.28

Table 5 displays participants' views regarding the effectiveness and washback of the testing and assessment process at EFL departments. It can be noted that the perception of the participants of Item 3: *test and assessment reflect course content* is positive with mean value 3.29 and  $p=0.00$  which is statistically significant indicating a strong evidence that they agree that the tests and assessments do reflect the course content they were studying. Neutral perceptions can be observed with the rest of the Items 1,2,4,5,6,7, and 8 as their p-values are greater than 0.05.

On the basis of the statistical analysis of data presented in Table 5, the testing and assessment process needs further improvement in terms of effectiveness and washback specifically in the areas, such as *tests and assessment result in better learning*, *test and assessment cover range of language skills*, *test and assessment help identify areas to focus on*, *tests and assessment influences the way students study*, *test and assessment motivate students to study effectively*, *test and assessment have positive impact on learning strategies*, and *test and assessment engage students in class activities*.

**Table (6)** Scoring and Grading (SG)

No.	item	N	Mean	Std. Deviation	t	p-value
1	Scoring is fair.	116	2.78	1.264	-1.837	0.07
2	Scoring is transparent.	116	3.09	1.092	0.850	0.40
3	Scores do not reflect performance.	116	3.25	1.243	2.166	0.03
4	Scoring criteria are applied to all students uniformly.	116	3.07	1.214	0.612	0.54
5	Grading criteria is clearly stated to students.	116	3.15	1.097	1.438	0.15
6	Grading process takes all students' skills into account.	116	3.00	1.142	0.000	1.00
7	Grades are provided in a timely manner.	116	3.22	1.135	2.127	0.04
8	Students understand how their final grade is calculated.	116	3.28	1.178	2.601	0.01

As shown in Table 6, participants' views of Item 1, *scoring is fair*, can be argued to be negative with mean value 2.78 and  $p=0.07$ . This suggests that Kurdish EFL students tend to disagree that scoring is fair in the testing and assessment process. Results in Table 6 also show that participants' perceptions of Item 3, *scores do not reflect performance*, is neutral with mean value 3.25 and  $p=0.03$  indicating that assessment at EFL departments is not totally inclusive in terms of employing different types of assessments and that there is a heavy focus on a particular type of assessment. Neutral perceptions can be observed in Table 6 with Items 2, 4, 5, and 6, *scoring is transparent*, *scoring criteria are applied to all students uniformly*, *grading criteria is clearly stated to students*, and *grading process takes all students' skills into account* as their p-values are greater than 0.05. This finding indicates that improvement is required in these areas, such as transparency, application of scoring process to all students in an equal manner, provision of grading process to students, and taking all

students skills into account when grading them. Also, neutral perceptions can be noted with items 7 and 8, *grades are provided in a timely manner* and *students understand how their final grade is calculated*, with mean values 3.22 and 3.28 and  $p= 0.04$  and  $0.01$  respectively. This also demonstrates that more efforts are needed for the grading process to be transparent in the provision of grades in the Kurdish EFL departments.

Feedback process in testing and assessment is crucial for it helps learners engage in learning. Table 7 displays participants' views of the feedback process conducted at EFL departments at public universities in the KRI.

**Table (7)** Feedback Process (F)

No.	item	N	Mean	Std. Deviation	t	p-value
1	Feedback is given in a timely manner.	116	2.91	1.116	-0.832	0.41
2	Feedback is supportive.	116	3.02	1.142	0.163	0.87
3	Feedback shows my weaknesses and strengths.	116	3.12	1.166	1.115	0.27
4	Feedback clarifies the grading criteria.	116	3.00	1.165	0.000	1.00
5	Feedback is constructive.	116	3.14	1.243	1.195	0.23
6	Students can discuss feedback with teachers.	116	3.01	1.205	0.077	0.94
7	Feedback is specific.	116	3.09	1.237	0.826	0.41
8	Feedback motivates better performance.	116	3.28	1.264	2.425	0.02

As it is clearly shown in Table 7, statistically significant difference can only be found with Item 8, *feedback motivates better performance*, with mean value 3.28 and  $p=0.02$ , indicating that there is an agreement by the participants that the feedback provided by teachers has a great role in enhancing better student performance. Strikingly, all the other items in the table received neutral responses implying no significant differences across all the items. This finding suggests that improvement is required in the following areas at EFL context at public universities in the KRI:

- Provision of a timely feedback.
- Provision of supportive feedback.
- Provision of feedback that pinpoints weak and strong areas in students' performance.
- Provision of clear and transparent feedback based on which grading is done.
- Provision of constructive feedback.
- Permitting students to discuss feedback with teachers.
- Provision of specific feedback for improvement purposes.

Is there statistically significant difference at the level of 0.05 between the mean values of the perceptions of students regarding the test and assessment process at EFL department according to the six variables analyzed above (D, AD, P, EW, SG, & F)? To verify this, the researchers

calculated the mean values and standard deviation for the data across all the above mentioned criteria (Table 8).

**Table (8)** Means Values

Field	N	Mean	Std. Deviation
D	116	3.0735	0.74107
AD	116	3.1380	0.75075
P	116	3.1313	0.69345
EW	116	3.1391	0.71517
SG	116	3.1080	0.69307
F	116	3.0743	0.86497
Total	696	3.1107	0.74323

After that, the researchers conducted a one-way analysis of variance (ANOVA) to compare the mean values across the six criteria: D, AD, P, EW, SG, and F. Descriptive statistics indicate that the mean values for the six criteria range from 3.0735 (F) to 3.1380 (AD), with standard deviations ranging from 0.69307 (SG) to 0.86497 (F). The total sample across the criteria is  $N=696$ .

**Table (9)** ANOVA result for Group Comparison

ANOVA					
Evaluation					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	0.544	5	0.109	0.196	0.964
Within Groups	383.373	690	0.556		
Total	383.917	695			

As indicated in Table 9, it is clear that the sig. value 0.964 is greater than 0.05, which indicates that there are no statistically significant differences across all the six criteria: D, AD, P, EW, SG, and F. The researchers attribute this to a number of reasons, such as testing and assessment strategies are applied uniformly to all students across all EFL department at the public universities in KRI, the testing and assessment regulations are centralized and carefully followed by the administration of the EFL departments, the design, administration, purpose, effectiveness and washback, feedback, and scoring and grading processes are conducted professionally by EFL teachers.

## 5. DISCUSSION OF RESULTS

The results arrived at in the present study show that the design of tests and assessments has been revealed to effectively been aligned with course objectives and the content of the tests and assessments were heavily related to the course content. This finding is in line with Pond's

(2019) report on the design of assessments using constructive alignment, stating that in the education system, if the tests and assessments align with the learning outcomes and objectives, a crucial foundation is then set for teaching, learning, and assessment. However, in terms of other areas pertinent to the design of tests and assessment, such as easiness of test and assessment instructions, difficulty level of question items, effective demonstration of learning through tests and assessments, the encouragement of critical thinking skills, and the coverage of the whole course material in testing and assessment, it has been revealed that all these areas are in need of improvement. French et al. (2023) agree that such areas are difficult to address especially when the test and assessment are conducted as high-stake evaluation. This is because if the tests and assessments are poorly designed, they might only encourage the memorization of course contents, thus leading students to focus on a particular bit of the material and ignore the rest. This has already been found out in this study as tests and assessment do not encourage critical thinking skills. This result aligns with the results of several other studies (Boud & Falchikov, 2006; Williams, 2014).

As for the administration of the testing and assessment process, on a holistic level, the findings revealed concerns about the quality of test and assessment administration. In terms of invigilation and test and assessment rules, it has been found out that the rules applied during invigilation do not negatively affect students' performance. This further implies that in terms of administration of assessment and invigilation, effective strategies and rules are employed by Kurdish EFL departments. This finding is in line with the results by Crossley (2022) and Van Bergen and Lane (2014), who state that effective invigilation and application of rules yields favored outcomes and minimizes opportunities for cheating. However, it also revealed that Kurdish EFL students were concerned about the misconduct of students in terms of cheating. Moreover, students expressed concerns about the quality of the seats provided as they disagreed that they were comfortable and did not cause distraction during the assessment process. This finding is in line with Cheryan et al. (2014) who state that the quality of seats and their design and arrangement play a significant role in students' achievement and engagement in the learning process.

Moreover, it is found out that the main purpose of testing and assessment is to check students' achievement and to pinpoint their weaknesses and strengths in terms of language learning. This suggests that the heavy focus of assessment and testing is on summative type as it checks students' achievement and provides feedback on their overall performance. This finding is supported by Smith and Fey (2000) who state that if an assessment is only directed to check students' final achievement, then it could be valid for some purposes such as provision of pass/fail decision and invalid or moderate at some other points such as predicting career achievement. It has also been found out that the purpose of testing and assessment is not to measure what students can do in terms of language ability, not to check students' progress, could be only for grading purposes, not to improve the pedagogical strategies in teaching, not for the purpose of checking students'

engagement in the activities and overall discussions. This confirms that testing and assessment at Kurdish EFL departments is only for summative purposes and decision-making on the final students' learning achievement. This further impacts the effectiveness and washback of the whole process as it has been confirmed that the testing and assessment process is not effective in terms of bringing about improved learning, covering a range of language skills, helping students identify what to focus on for improvement purposes, encouraging students to engage in effective study practices, fostering beneficial effect on learning strategies, and actively involving students in classroom activities. This finding is in harmony with Benediktsson and Ragnarsdóttir (2020), Gijbels and Dochy (2006), and Wang and Brown (2014) who consider such testing and assessment process as demotivating or has no positive impact on the students' learning and achievement as students' prefer to have assessment strategies that could leave impact on their learning and involve them deeper in the process of learning. Villarroel et al. (2019) agree that testing and assessment process which is ineffective does not lead to a favored washback/backwash. Tests and assessments, as they state, should be designed to support higher levels of thinking and critical involvement.

In line with the above results, it has been found out that scoring students' answers is unfair and that the grades they obtain do not represent their actual performance. This confirms that testing and assessment in Kurdish EFL context at HE public universities is heavily depended on summative type of assessment. This is confirmed by Knight (2002) who stated that summative type of assessments does not provide learners with the opportunity to improve their performance and learn from their mistakes. Moreover, the grading process has been found not to represent all students' skills and that the scoring criteria are not transparent and applied consistently to all students. These findings are in line with the findings by Salehi et al. (2019) who found out that scoring and grading process could generate academic inequity due to heavy dependence on specific skills and ignorance of others. However, the positive findings about the testing and assessment process is that grades are given to students in a timely manner and that students are aware of the way their final grades are calculated based on their achievement during the course.

In terms of feedback in the testing and assessment process, it has been found out that for an improved process of assessment, further work needs to be done in terms of feedback provided to students based on their performance. This result is supported by Winstone and Carless (2020) who stated that the provision of a quality feedback in a timely manner boosts students' overall performance and helps them engage with the learning process and makes sure performance is enhanced. However, it can be deduced that since the nature of the testing and assessment process is summative, quality feedback that supports ongoing learning and active engagement of students could be missing as this is more related to formative assessment (Henderson et al. 2020).

Overall, it can be argued that the results of the current study are aligned with national and international studies

referred to in this study. On a national level, the findings are consistent with those of Mahmood and Galeb (2024) who stated that the testing and assessment process utilized needs improvement on a number of levels and that the majority of the students are not satisfied with the ways they are assessed. Moreover, the assessment process in the EFL context of HE public universities needs to incorporate ongoing assessment methods. On an international level, the current study is in harmony with Vavla and Gokaj (2013) and Woolever (2019) who found out that students viewed assessment as unfair, do not yield effective learning, and are not effective measure in determining students' overall language potentials.

## 6. CONCLUSION

Based on the above findings and discussion, the current research has arrived at the following concluding points:

1. There is an effective alignment of testing and assessment items with course content and course objectives. Also, the items of the questions and the overall assessment instructions are difficult for students; tests and other assessments do not demonstrate effective learning, tests and assessment are designed in a way which does not involve critical thinking skills, and the materials and skills covered are not all included in the testing and assessment process.
2. As a process, effective invigilation and testing and assessment rules are applied in a way that minimizes misconduct. However, concerns about students' cheating and unsupportive environment have been other issues that were found out. Moreover, the poor condition of seats and halls negatively impacts students' performance on the tests and during other assessment processes.
3. Regarding the purpose of testing and assessment, it has been found out that the only purpose of testing and assessment is to measure students' achievement and showcase their weak and strong skills, hence ignoring other essential purposes of testing and assessment, such as checking the overall language ability of the students, enhancing teaching by tracing students' progress, facilitating students' engagement in the learning process, and avoiding the use of tests and assessment process for grading purposes only.
4. The process of testing and assessment is inefficient in demonstrating students' overall learning and that a comprehensive coverage of a range of language skills is missing.
5. There are doubts about whether or not scoring is fair and does not represent the overall performance of the students. However, the provision of students' grades and the way they are calculated are proved to be effective.
6. The testing and assessment process lacks effective, supportive, and constructive feedback on students' overall performance and work.

## 7. REFERENCES

- Al-Wadi, H. (2020). Bahrain's secondary EFL teachers' beliefs of English language national examination: 'How it made teaching different?'. *International Journal of Instruction*, 13(1), 197–214. <https://doi.org/10.29333/iji.2020.13113a>
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115–129.
- Aslam, R., & Khan, N. (2020). Constructive feedback and students' academic achievement: A theoretical framework. *New Horizons*, 14(2), 175–198. [https://doi.org/10.2.9270/NH.14.2\(20\).10](https://doi.org/10.2.9270/NH.14.2(20).10)
- Bachman, L. F., & Palmer, A. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F. and Palmer, A. S. (2010). *Language Assessment in Practice*, 2nd edn. Oxford: Oxford University Press.
- Bachman, L., & Damböck, B. (2018). *Language assessment for classroom teachers*. Oxford University Press.
- Babbie, E. (2021). *The practice of social research* (16th ed.). Wadsworth Publishing.
- Basera, C. H. (2019). Learners' perceptions of assessment strategies in higher education. *Journal of Education and e-Learning Research*, 6(2), 76–81. <https://doi.org/10.20448/journal.509.2019.62.76.81>
- Benediktsson, A. I., & Ragnarsdóttir, H. (2020). Immigrant students' experiences of assessment methods used in Icelandic universities. *Multicultural Education Review*, 12(2), 98–116. <https://doi.org/10.1080/2005615X.2020.1756090>
- Bennett, R. E. (2011). Formative assessment: A critical review. *Assessment in Education: Principles, Policy & Practice*, 18(A), 5–25.
- Biggs, J., & Tang, C. (2011). *Teaching for quality learning at university: What the student does* (4th ed.). Open University Press.
- Black, P. J. (1993). Formative and summative assessment by teachers. *Studies in Science Education*, 21(1), 49–97.
- Black, P., & Wiliam, D. (1998). "Assessment and classroom learning." *Assessment in Education: Principles, Policy & Practice*, 5(1), 7–74.
- Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability*, 21(1), 5–31.
- Black, P., & Wiliam, D. (2018). Classroom assessment and pedagogy. *Assessment in Education: Principles, Policy & Practice*, 25, 551–575. <https://doi.org/10.1080/0969594X.2018.1441807>
- Boud, D., & Falchikov, N. (2006). Aligning assessment with long-term learning. *Assessment and Evaluation in Higher Education*, 31(4), 399–4
- Brace, I. (2008). *Questionnaire Design: How To Plan, Structure And Write Survey Material for Effective Market Research*. Kogan Page.
- Broadfoot, P. (2007). *An introduction to assessment*. Continuum International Publishing Group.
- Brookhart, S. M. (2008). *How to provide effective feedback to students*. ASCD, 1–12.
- Brookhart, S. M. (2013). *Grading and learning: Practices that support student achievement*. ASCD.



- Brookhart, S. M. (2017). *How to use grading to improve learning*. ASCD.
- Brookhart, S. M., & Nitko, A. J. (2019). *Educational assessment of students* (8th ed.). Pearson.
- Brown, D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices* (3rd ed.). Pearson Longman, New York.
- Brown, G. T. L., & Harris, L. R. (2016). The future of assessment research as a human and social endeavour. In G. T. L. Brown & L. R. Harris (Eds.), *Handbook of Human and Social Factors in Assessment* (pp. 506–523). Routledge.
- Brown, H. D., & Abeywickrama, P. (2019). *Language assessment: Principles and classroom practices*. Pearson.
- Butler, Y. G. (2021). "Assessing young learners," in The Routledge Handbook of Language Testing. 2nd Edn. eds. G. Fulcher and L. Harding (New York: Routledge), 153–170.
- Carless, D. (2015). *Excellence in university assessment: Learning from award-winning practice*. Routledge.
- Cheng, L. (2005). *Changing language teaching through language testing: A washback study*. Cambridge University Press.
- Cheng, L. (2017). Washback in language testing: Research contexts and methods. *Language Testing*, 34(4), 473–481. <https://doi.org/10.1177/0265532217710658>
- Cheng, L., & Green, D. W. (2007). The washback effect of EFL standardized tests in China: A case study. *Language Testing*, 24(3), 321–346.
- Cheryan, S., Ziegler, S. A., Plaut, V. C., & Meltzoff, A. N. (2014). Designing classrooms to maximize student achievement. *Policy Insights from the Behavioral and Brain Sciences*, 1(1), 4–12. <https://doi.org/10.1177/2372732214548677>
- Cooper, D. R., & Schindler, P. S. (2014). *Business research methods* (12th ed.). McGraw-Hill Education.
- Crossley, M. (2022). Merlin Crossley makes the case for exams. *Campus Morning Mail*. <https://campusmorningmail.com.au/news/merlin-crossley-makes-the-case-for-exams/>. Accessed November 29<sup>th</sup>, 2024
- Dabiri, A. (2018). A critical discourse analysis on teachers' verbal feedback patterns in EFL CLT classrooms. *Journal of English Educators Society*, 3(2), 129. <https://doi.org/10.21070/jees.v3i2.1262>
- Dixson, D. D., & Worrell, F. C. (2016). Formative and summative assessment in the classroom. *Theory Into Practice*, 55(2), 153–159.
- Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.
- Fitriyah, I., Bastomi, Y., Khotimah, K. & Gozali, I. (2022). Implementation of Assessment for Learning in Online EFL Writing Class: A Case of Novice Undergraduate Teachers. *LEARN Journal: Language Education and Acquisition Research Network*, 15(2), 129-159.
- French, S., Dickerson, A., & Mulder, R. A. (2024). A review of the benefits and drawbacks of high-stakes final examinations in higher education. *Higher Education*, 88(6), 893–918. <https://doi.org/10.1007/s10734-023-01148-z>
- Frey, B. B., & Schmitt, V. L. (2007). Coming to terms with classroom assessment. *Journal of Advanced Academics*, 18, 402–423.
- Fulcher, G. (2010). *Practical Language Testing*. Hodder Education/Routledge. <https://doi.org/10.4324/980203767399>
- Fulcher, G. (2013). Scoring performance tests. In G. Fulcher & F. Davidson (Eds.), *The Routledge handbook of language testing* (pp. 392–406). Routledge. <https://doi.org/10.4324/9780203181287>
- Fulcher, G. (2019). Cultivating language assessment literacy as collaborative CPD. In M. Gillway (Ed.), *Addressing the State of the Union: Working Together, Learning Together* (pp. 27–35). Garnet.
- Fulcher, G., & Davidson, F. (2007). *Language testing and assessment: An advanced resource book*. Routledge.
- Gijbels, D., & Dochy, F. (2006). Students' assessment preferences and approaches to learning: Can formative assessment make a difference? *Educational Studies*, 32(4), 399–409. <https://doi.org/10.1080/03055690600850354>
- Grabin, L. A. (2009). *Alternative assessment in the teaching of English as a foreign language in Israel* (Unpublished doctoral dissertation). Pretoria, Gauteng, South Africa.
- Guskey, T. R., & Brookhart, S. M. (2019). *What We Know About Grading: What Works, What Doesn't, and What's Next*. ASCD.
- Harlen, W. (2021). *Assessment and learning: State of the art*. SAGE Publications.
- Hattie, J., & Clarke, S. (2019). *Visible learning: Feedback*. Routledge.
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112.
- Henderson, M., Ajjawi, R., Boud, D., & Molloy, E. (Eds.). (2020). *The impact of feedback in higher education: Improving assessment outcomes for learners*. Palgrave Macmillan.
- Hughes, A. (2003). *Testing for language teachers* (2nd ed.). Cambridge University Press.
- Hultgren, A. K., N. Owen, P. Shrestha, M. Kuteeva and Š. Mežek. 2022. Assessment and English as a medium of instruction: Challenges and opportunities. *Journal of English-Medium Instruction* 1(1): 105 – 123. DOI: <https://doi.org/10.1075/jemi.21019.hul>
- Jerrim, J. (2022). Test anxiety: Is it associated with performance in high-stakes examinations? *Oxford Review of Education*. <https://doi.org/10.1080/03054985.2022.2079616>
- Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity, and educational

- consequences. *Educational Research Review*, 2(2), 130-144.
- Joshi, A., Kale, S., Chandel, S. and Pal, D. (2015) Likert Scale: Explored and Explained. *British Journal of Applied Science & Technology*, 7, 396-403. <https://doi.org/10.9734/BJAST/2015/14975>
- Kellaghan, T., & Greaney, V. (2003). *Using assessment to support student learning*. UNESCO Publishing.
- Klatte, M., Bergström, K., & Lachmann, T. (2013). Does noise affect learning? A short review on noise effects on cognitive performance in children. *Frontiers in Psychology*, 4, 578. <https://doi.org/10.3389/fpsyg.2013.00578>
- Knight, P. T. (2002). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education*, 27(3), 275–286. <https://doi.org/10.1080/03075070220000662>
- Koh, J. H. L. (2014). How does formative assessment influence students' self-regulation? *Asia Pacific Education Review*, 15(2), 195–206.
- Lado, R. (1961). *Language Testing: The Construction and Use of Foreign Language Tests*. A Teacher's Book. McGraw-Hill Book Company.
- Lin, J., Li, T., Tsai, Y.-S., Gasevic, D., & others. (2023). Can large language models provide feedback to students? A case study on ChatGPT. *Preprint*. <https://doi.org/10.35542/osf.io/hcgzi>
- Lynch, B. K. (2001). *The assessment of language for specific purposes*. Cambridge University Press.
- Lynch, B. K. (2003). *Language assessment and teaching: A critical perspective*. Routledge.
- Mahmood, S. K., & Ghaleb, N. (2024). Kurdish EFL students' perceptions towards summative and formative assessment at Salahaddin University. *Journal of Tikrit University for the Humanities*, 31(5), Article 25. <https://doi.org/10.25130/jtuh.31.5.2024.25>
- Manitoba Education, Citizenship and Youth, (2006). *Rethinking Classroom Assessment with Purpose in Mind: assessment for learning, assessment as learning, assessment of learning*.
- McMillan, J. H. (2011). *Classroom assessment: Principles and practice for effective standards-based instruction*. Allyn & Bacon.
- Mertens, D. M. (2019). *Research and Evaluation in Education and Psychology: Integrating Diversity with Quantitative, Qualitative, and Mixed Methods*. Sage Publications.
- Messick, S. (1996). Validity and washback in language testing. *Language Testing*, 13(3), 241-256. <https://doi.org/10.1177/026553229601300302>
- Mngomezulu, H., Ramaila, S., & Dhurumraj, T. (2022). Pedagogical strategies used to enact formative assessment in science classrooms: Physical sciences teachers' perspectives. *International Journal of Higher Education*, 11(3), 158–166. <https://doi.org/10.5430/ijhe.v11n3p158>
- Morrow, K. (1986). The evaluation of tests of communicative performance. In M. Portal (Ed.), *Innovations in language testing* (pp. 1-13). National Extension College Trust.
- Moss, P. A., Girard, B. J., & Haniford, L. C. (2006). Validity in educational assessment. *Review of Research in Education*, 30(Special Issue: Rethinking Learning: What Counts as Learning and What Learning Counts), 109–162. American Educational Research Association. <http://www.jstor.org/stable/4129771>
- Murad, I. (2015). The effects of Kurdish learners' characteristics on their English language learning. *European scientific journal*, 11(10).
- Murphy, D. H., Little, J. L., & Bjork, E. L. (2023). The value of using tests in education as tools for learning—not just for assessment. *Educational Psychology Review*, 35, 89. <https://doi.org/10.1007/s10648-023-09808-3>
- Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of students* (6th ed.). Pearson.
- Obilor, E. I. (2019). Feedback and students' learning. *International Journal of Innovative Education Research*, 7(2), 40–47.
- Paechter, M., Maier, B., & Macher, D. (2010). Students' expectations of, and experiences in e-learning: Their relation to learning achievements and course satisfaction. *Computers & Education*, 54(1), 222–229. <https://doi.org/10.1016/j.compedu.2009.08.005>
- Pond, K. (2019). *University education: Assessment and assessment design*. School of Business and Economics, Loughborough University.
- Popham, W. J. (2008). *Transformative assessment*. ASCD.
- Popham, W. J. (2013). *Classroom assessment: What teachers need to know* (7th ed.). Pearson.
- Putri, N. S. E., Pratolo, B. W., & Setiani, F. (2019). The alternative assessment of EFL students' oral competence: Practices and constraints. *Ethical Lingua: Journal of Language Teaching and Literature*, 6(2), 72–85. <https://doi.org/10.30605/25409190.v6.72-85>
- Qadir, S. M., Omar, R. M., Rasheed, M. H., & Mohammed, C. J. (2023). Assessing the end-of-semester examination papers during the implementation of the Bologna Process: Bloom's Taxonomy as a framework. *Koya University Journal of Humanities and Social Sciences*, 6(1), 77–87. <https://doi.org/10.14500/kujhss.v6n1y2023.pp77-87>
- Sadler, D. R. (2009). Grade integrity and the representation of academic achievement. *Studies in Higher Education*, 34(7), 807-826.
- Sauro, J., & Lewis, J. R. (2016). *Quantifying the user experience: Practical statistics for user research* (2nd ed.). Morgan Kaufmann.
- Sekaran, U., and Bougie, R. (2020). *Research methods for business: A skill-building approach* (8th ed.). Wiley.



- Smith, M. L., & Fey, P. (2000). Validity and accountability in high-stakes testing. *Journal of Teacher Education*, 51(5), 334–344. <https://doi.org/10.1177/0022487100051005002>
- Stiggins, R. J., & Chappuis, J. (2017). *An introduction to student-involved assessment FOR learning* (7th ed.). Columbus, OH: Pearson Education.
- Salehi, S., Cotner, S., Azarin, S. M., Carlson, E. E., Driessen, M., Ferry, V. E., Harcombe, W., McGaugh, S., Wassenberg, D., Yonas, A., & Ballen, C. J. (2019). Gender performance gaps across different assessment methods and the underlying mechanisms: The case of incoming preparation and test anxiety. *Frontiers in Education*, 4. <https://doi.org/10.3389/educ.2019.00107>
- Taras, M. (2005). Assessment—summative and formative—some theoretical reflections. *British Journal of Educational Studies*, 53(4), 466–478.
- Tierney, R. D. (2014). Fairness in educational assessment. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 131–148). Springer.
- Tsagari, D., & Vogt, K. (Eds.). (2017). *Handbook of assessment for language teachers*. TALE Erasmus+ Project. <https://www.taleproject.eu>
- Vavla, L., & Gokaj, R. (2013). Learner's perceptions of assessment and testing in EFL classrooms in Albania. *Mediterranean Journal of Social Sciences*, 4(11), 509. MCSER Publishing.
- Van Bergen, P., & Lane, R. (2014). Exams might be stressful, but they improve learning. *The Conversation*. <https://theconversation.com/exams-might-be-stressful-but-they-improve-learning-35614>. Accessed 11/29/2024
- Villarroel, V., Boud, D., Bloxham, S., Bruna, D., & Bruna, C. (2019). Using principles of authentic assessment to redesign written examinations and tests. *Innovations in Education and Teaching International*, 1–12. <https://doi.org/10.1080/14703297.2018.1564882>
- Voinea, L. (2018). Formative assessment as assessment for learning development. *Revista de Pedagogie – Journal of Pedagogy*, 66(1), 7–23. <https://doi.org/10.26755/RevPed/2018.1/7>
- Wang, C., Hancock, D., Shieh, J.-J., & Hachen, J. (2023). Student perceptions of assessment in Taiwan and the United States. *Educational Research and Development Journal*, 26(2), 62–84.
- Weir, C. J. (1990). *Communicative language testing*. Prentice Hall.
- Weir, C. J. (2005). *Developing language tests*. Cambridge University Press.
- William, D. (2018). “Feedback: At the Heart Of—But Definitely Not All of—Formative Assessment,” in *Cambridge Handbook of Instructional Feedback*. Editors A. A. Lipnevich, and J. K. Smith (Cambridge: Cambridge University Press), 376–408.
- Williams, P. (2014). Squaring the circle: A new alternative to alternative-assessment. *Teaching in Higher Education*, 19(5), 565–577. <https://doi.org/10.1080/13562517.2014.882894>
- Winstone, N. E., & Carless, D. (2020). *Designing effective feedback processes in higher education: A learning-focused approach*. Routledge, Taylor & Francis Group.
- Woolever, J. L. (2019). *Student perceptions towards mandated assessments* (Doctor of Education dissertation). Educational Leadership & Policy Studies, University of Kansas.

## فەكۆلنەك ل سەر بۆچوونىن خویندكارىن كورد بىن كۆ زمانى نىنگلىزى وەك زمانەكى بىانى دخوینن، دەربارەى سیستەمى ھەلسەنگاندنى داناستى زانیگەھى دا

پوختە:

نارمانج ژ ئەقئ فەكۆلنەك؛ فەكۆلنەك ل سەر بۆچوونىن خویندكارىن كورد، ئەوین فیری زمانى نىنگلىزى دىن وەك زمانەكى بىانى ل دامودەزگەھىن خواندنا بلند و گشتى ل ھەرمە كوردستانا عێراق. سیستەمى ھەلسەنگاندنى نىكە ژ بنگەھىن سەرەكى د فیربوون و پېشقەچوونا خویندكاران دا، ھەروەسا كارىگەرەكا مەزن یا ھەى د پشكداریكرنا وان د پروسيسا فیربوونى دا، پېدقېه گرنگیەكا تابیەت پى بېتەدان. نارمانجا سەرەكىا قى فەكۆلنەك ئەو فەكۆلنەك ل سەر بۆچوونىن خویندكارىن كورد ل سەر (چېكرن، برێقەبرن، مەرم، كوالیتى، كارىگەرى) ھەروەسا شیوازى نرەخاندنا بەرسقین خویندكاران د سیستەمى ھەلسەنگاندنا وان دا. بۆ قى مەرمە پرسینامەكا تابیەت بۆ كۆمكرنا پیزانىنان دەربارەى بۆچوونىن خویندكاران ھاتیە ئامادەكرن و برێكا لىكئى گۆگل فورم، بۆ خویندكارىن پشكین زمانى نىنگلىزى بىن سیمستەرىن (3,5,7) بۆ سالا خواندنى (2024-2025) ل ھژمارەك ژ زانكۆیین حكومى ھاتیە ھنارتن. بۆ شروەكرنا باوەریا خالین راپرسی (كرونباخ الفا Cronbach Alpha) ھاتیە بكارینان و بۆ شروەكرنا تىكرایا ژمیریاری پروگرامى SPSS گۆھارتویا 25 ھاتیە بكارینان. ھەروەسا بۆ دیتنا جوداھیا تىكرایا ژمیریاری دناقیەر ھەر شەش تىگۆھەران دا. مە پروگرامى ANOVA بكاریناپە. ئەنجامین فەكۆلنەك وەسا دیاركرن: كۆ گەلەك ناستەك بىن ھەین د پروسيسا بېھننا سیستەمى ھەلسەنگاندنى ب شیوازەكى دروست ل دامودەزگەھىن خواندنا بلند. ژ لایەكى دیقە ل دەمى كۆ كەرسەیین ھەلسەنگاندنى لگەل نارمانجین پەرومردەن د ھەماھەنگن، بەلى گەلەك جاران ئەق كەرسە و شیوازى دروستكرنا وان ژ لایە مامۆستایان قەب دروستى ھەرا بەرفرەو و رەخنەى ناھینە پێقان. ژ لایە پروسيسین كارگىرى و نەچاركرنا نەدیار یان قەدەغەیان و سەرراستكرن و نرەخاندنا نەدرست و نە یەكسان لگەل سەرەنچا كیشیا زیدە ل سەر تاقىكرن نىن نفیسكى ل جەھى كۆ پېشكەفتن و تىكەلبوونا خویندكاران زیدە بكەت، كار تىكرنا نرەخاندنى بەمەز تر دكەت. فاكترەین ژینگەھا ئەزموونان وەك دینامىكین ژوورین خواندنى نە د گونجایی نە، كوالیتیا كۆرسىكین خویندكاران بىن نەباش كار تىكرنەكا نەرىنى ل سەر شیواز و كوالیتیا بەرسقاندنا وان دكەت. ھەروەسا كۆمبوونا فیدبەكەكا ئافاكەر كار تىكرنەك نەرىنى ل سەر ئەنجامین زانستیین وان دكەت.

پەقیق سەرەكى: بۆچوونىن خویندكاران، چېكرنا پسارن، برێقەبرنا تاقىكرن، زمانى نىنگلىزى وەك زمانەكى بىانى، فیدبەك، كوالیتیا پسارن.

## التقصي عن آراء الطلبة الكرد الذين يدرسون اللغة الإنجليزية كلغة أجنبية حول عملية التقييم في المستوى الجامعي

الخلاصة:

تهدف الدراسة الحالية إلى تقصي عن آراء طلبة اللغة الإنجليزية حول عملية التقييم التي يتم إجراؤها في أقسام اللغة الإنجليزية كلغة أجنبية في الجامعات الحكومية في إقليم كردستان العراق. لأن التقييم هو العامل الأساسي لتعلم الطلبة ومشاركتهم وتقييمهم بالإضافة إلى المقياس الوحيد لتقدمهم وتطورهم، يجب إيلاء الكثير من الاهتمام لعملية التقييم. تهدف هذه الدراسة على وجه التحديد إلى دراسة اتجاهات الطلبة الكرد الذين يتعلمون اللغة الإنجليزية كلغة أجنبية تجاه تصميم وإدارة وفعالية الاختبارات والغرض منها وتأثيرها على العملية التعليمية. وكذلك التقصي عن عملية التصحيح والتغذية الراجعة تجاه عملية الاختبار والتقييم. ولإجراء هذه الدراسة تم استخدام استبانة صممت خصيصاً لجمع البيانات المطلوبة للدراسة من (116) طالباً في الفصول الدراسية 3 و 5 و 7 والتي تمثل المراحل الجامعية (الثانية والثالثة والرابعة) في أقسام اللغة الإنجليزية في عدد من الجامعات الحكومية في إقليم كردستان العراق خلال العام الدراسي (2024-2025). ولغرض التحقق من موثوقية عناصر الاستبانة، تم استخدام معامل كرونباخ ألفا جنباً إلى جنب مع برنامج (SPSS إصدار 25) لتحليل متوسط قيم العناصر وتم استخدام تحليل التباين (ANOVA) لمقارنة متوسط القيم عبر المجموعات الست. أشارت النتائج إلى تحديات كبيرة في موازنة وتنفيذ عمليات الاختبار والتقييم في المستوى الجامعي. في حين أن عناصر الاختبار والتقييم تتوافق مع الأهداف التعليمية، إلا أنها غالباً ما تفشل في قياس التفكير النقدي ومهارات اللغة الشاملة بشكل كافٍ. وأدت الإجراءات الإدارية والتي تتضمن إصدار تعليمات غير واضحة للطلبة حول الامتحانات، وانعدام الشفافية والموضوعية أثناء عملية تصحيح الدفاتر الامتحانية، والتركيز المفرط على الامتحانات التحريرية بدلاً من تعزيز مهارات الطلبة اللغوية والأكاديمية، وعلاوة على ذلك، تؤثر العوامل البيئية مثل الغش، وعدم توفر القاعات الامتحانية الملائمة، وندرة المقاعد الدراسية سلباً على أداء الطلبة. وكما أن الافتقار إلى التغذية الراجعة البناءة تؤدي إلى عدم تطوير المهارات العامة للطلبة. وتؤكد هذه الدراسة على الحاجة الملحة في تبني برنامج شامل للتقييم يساهم في نمو مهارات الطالب وجعل عملية التقييم أكثر شفافية وموضوعية بهدف دمج الكفاءات اللغوية المتنوعة.

الكلمات المفتاح: آراء الطلبة الكرد، تصميم وإدارة اختبارات، اللغة الإنجليزية كلغة أجنبية، عملية التصحيح، التغذية الراجعة، فعالية التقييم.