Arab World English Journal
INTERNATIONAL PEER REVIEWED JOURNAL ISSN: 2229-9327

Arab World English Journal (AWEJ) Volume 16. Number 2, June 2025 Pp. 373-393
DOI: https://dx.doi.org/10.24093/awej/vol16no2.21

Exploring Students' Perceptions of Assessment and Testing Practices and Their Impact on Learning Outcomes in Moroccan Higher Education

Anouar Mohamed EL KASRI
School of Arts and Humanities, Moulay Ismail University, Meknes, Morocco
Corresponding Author: Anouar.elkasri@yahoo.com

Moulay Sadik MALIKI
School of Arts and Humanities, Hassan II University of Ain Chock, Casablanca, Morocco
Email: sadikmaliki@yahoo.fr

Mohammed LAROUZ
School of Arts and Humanities, Moulay Ismail University of Meknes, Morocco
Email: mlarouz@yahoo.com

Martina PAVLÍKOVÁ
Department of Journalism and New Media, Faculty of Arts, Constantine the Philosopher University in Nitra, Slovakia
Email: mpavlikova@ukf.sk

Brahim EL YOUSFI
Former Scientist at Plant Pathology Lab, INRA-CRRA, Morocco
Email: elyousfi_brahim@yahoo.com

Received: 02/24/2025 Accepted:06/19/2025 Published:06/26/2025

Abstract
Education research emphasizes the significance of students' perceptions of assessment. This study aimed to investigate how students perceive the assessment tasks designed by their professors and to examine the relationship between these perceptions and their learning outcomes (i.e., grades) across three Moroccan Faculties of Letters and Humanities. The study is particularly important as it provides valuable feedback to test designers within these faculties and supports efforts to improve the quality of assessments for common core students. Furthermore, it aims to raise professors' awareness of the impact that test quality has on student performance and encourage more thoughtful and research-informed approaches to assessment design. In this regard, the study answers the following research questions: How do Moroccan students in the departments of English studies at the faculties of letters and humanities perceive the assessment tasks designed by their professors concerning the validity, reliability, authenticity, washback, and practicality of assessment tasks? Is there a relationship between students' perception of assessment and their performance? To investigate the five variables associated with students' perceptions of assessment, 526 higher education students completed the adapted Students' Perception of Assessment Questionnaire (SPAQ). The obtained data were analyzed using descriptive statistics to demonstrate the overall perception of students. In addition, multiple regression of the five scales on students' grades was used to measure the relationship between students' perception and performance. The results revealed that students across the three Faculties had a positive perception of assessment. However, no significant relationship was found between students' perceptions of assessment and academic performance.

Keywords: Assessment criteria, learning outcomes, Moroccan higher education, perception of assessment, students' performance

How to cite: EL KASRI, A. M., MALIKI, M. S., LAROUZ, M., PAVLÍKOVÁ, M., & EL YOUSFI, B. (2025). Exploring Students' Perceptions of Assessment and Testing Practices and their Impact on Learning Outcomes in Moroccan Higher Education. *Arab World English Journal, 16*(2):373-393. https://dx.doi.org/10.24093/awej/vol16no2.21

Introduction

Classroom assessment is a key component of the teaching and learning process. Many researchers have explored how classroom assessment contributes to student learning (Harlen & Crick, 2003). In this context, the assessment tasks assigned in the classroom convey significant messages to students about their value, importance, and relevance (Alkharusi et al., 2014). This perspective arose from the need for assessment practices to align with 21st-century requirements (Othman & Bouyakoub, 2020). It was also the result of a shift toward constructivist learning approaches (Van de Watering et al., 2008) and the integration of innovative learning environments, including co-regulation techniques and their influence on learning outcomes (Ajmal et al., 2024). Moreover, it is a reaction to traditional approaches, prioritizing rote learning and passive memorization of grammatical rules and language items. In this context, Beg et al. (2024) investigated the integration of the Rotation Model into the ESL curriculum. They found that it significantly enhances language learning outcomes by employing active, dynamic instructional strategies that effectively manage cognitive load. Given the profound impact of assessment on learning, the effectiveness of educational programs largely depends on the quality of assessment practices (Van der Vleuten et al., 2017). High-quality assessment encompasses various aspects of evaluation, including language assessment criteria (Leeuwenkamp et al., 2019). When students understand the advantages of quality assessment and the negative consequences of poor assessment, they develop their perceptions of its value (Leeuwenkamp, 2019).

Students' perceptions of assessment hold great importance for several reasons. As the primary source of information for evaluating assessment methods, students' attitudes and viewpoints provide valuable insight into the quality of classroom assessment (Rahman, 2020). Their active participation in assessment also makes the learning process more engaging and meaningful (Rahman, 2020). According to Dhindsa et al. (2007), examining students' perceptions encourages teachers to adopt more authentic and realistic evaluation approaches in the future. Furthermore, students' perceptions of assessment influence their learning and impact their overall success in the classroom (Mussawy, 2009). Accordingly, educational institutions should integrate students' perceptions into assessment design to ensure quality practices that foster improved learning outcomes (Leeuwenkamp, 2019).

Despite the growing body of international research on student perceptions of assessment, the specific ways in which students perceive assessment tasks—and how these perceptions influence learning outcomes such as academic performance or grades—remain insufficiently explored. Thus far, this relationship has not been specifically examined within the Moroccan context. This is a notable gap, as Moroccan higher education presents a particularly compelling case for investigation. The system is undergoing significant educational reforms and characterized by a diverse linguistic and cultural landscape. Moreover, it is increasingly shifting toward more student-centered approaches in English language teaching and learning. These unique contextual factors may shape how assessment practices are perceived and experienced by students, which calls for the need to study this setting on its terms rather than relying solely on findings from other

educational systems. Therefore, the primary aim of this study was to address the identified gap by exploring how students in Moroccan higher education perceive assessment tasks and how these perceptions relate to their learning outcomes.

This study holds considerable significance for two key reasons. First, it offers test designers in Moroccan faculties of Letters and Humanities critical insights and constructive feedback regarding the quality and effectiveness of the assessments they develop for common core students. By evaluating these tests, the study helps identify strengths and weaknesses in their design, thereby guiding improvements that can lead to more valid, reliable, and fair assessments. Second, the study heightens university professors' awareness of how the quality of assessments directly influences student performance. It emphasizes the pivotal role that well-constructed tests play in accurately measuring students' knowledge and in shaping their academic experience. Drawing attention to these effects, the study encourages professors to adopt more reflective and research-based approaches to test design. To reach this objective, the study seeks to answer the following research questions :

RQ 1 - How do Moroccan students in the departments of English studies at the faculties of letters and humanities perceive the assessment tasks designed by their professors concerning the validity, reliability, authenticity, washback, and practicality of assessment tasks?

RQ 2- Is there a relationship between students' perception of assessment and their performance?

Literature Review

The Relationship between assessment and learning

The relationship between assessment and learning can be understood by examining the evolution of language testing and teaching methodologies over the past century. Historically, language instruction has been shaped by developments in linguistics, psychology, and second language acquisition (Richards & Rogers, 2016). In this context, language-testing practices have evolved alongside shifts in teaching methods (Brown & Abeywickrama, 2010). Traditional testing methods have been shaped by older paradigms like behaviorist learning theory, which treated testing as separate from instruction and relied heavily on formats such as multiple-choice and open-ended questions (Shepard, 2000).

In recent decades, the role of assessment in learning has shifted due to adopting a more constructivist learning paradigm, leading to the development of new learning environments (Van de Watering, 2008). The most significant change in the perception of assessment is the concept of using it as a tool for learning (Dochy & McDowell, 1997). Rather than primarily serving to assign grades and measure students' achievement of predefined objectives, there is now broad recognition that assessment plays a vital role throughout the entire learning process (Brown et al., 1994; Gibbs, 1999). Assessment influences learning for several reasons. First, it offers students valuable insights into their learning progress, goals, and strategies to achieve them (Watkins et al., 2005). In addition, how students perceive assessment influences how they prepare for it, which can positively or negatively impact their learning outcomes. (Nevo, 1995; Gielen et al., 2003).

Students' perception of assessment

In today's information-driven world, teachers must help their students take an active role in managing and monitoring their learning. This shift toward a more learner-centered approach calls for a fundamental change in students' roles within the teaching and learning process. To align with students' evolving role as "partners" in learning (Stiggins, 2001), teachers must adopt a perspective that takes into consideration the student's viewpoint (Horwitz, 1989). Likewise, curriculum designers should consider students' experiences and perspectives on their learning environment (Ekbatani, 2000; Lindblom-Ylanne & Lonka, 2001). In language testing and assessment, students are also recognized as key stakeholders (Ericson & Gustafsson, 2005). As a result, research has increasingly focused on students' perceptions of assessment, revealing that assessment tasks in the classroom send powerful messages about their importance (Black & Williams, 1998; Linnenbrink & Pintrich, 2001, 2002; McMillan & Workman, 1998). For example, a study by Dhindsa et al. (2007) on Bruneian secondary school students' perceptions of assessment tasks found that students generally viewed these tasks as transparent and aligned with classroom teaching and objectives despite reporting low levels of consultation, authenticity, and diversity in these assessment tasks.

Similarly, Rahman (2020) explored students' perceptions of grammar assessment at the English department of the State Islamic University of Ar-Raniry. Sixty-nine fourth-year English students participated in the study by completing the Students' Perception of Assessment Questionnaire (SPAQ), which evaluated the alignment of grammar assessment with planned learning, authenticity, transparency, and students' classroom capabilities. The findings indicated that students perceived only a slight alignment between grammar assessment and the intended learning outcomes. They also noted a lack of transparency regarding the purpose, authenticity, and forms of assessment. Additionally, the results revealed that students' performance in taking the assessment was generally unsatisfactory.

In a study involving 198 tenth-grade students in Oman, Alkharusi (2013) discovered that a high level of authenticity and transparency in assessment was positively associated with students' self-efficacy and perception of task value. Similarly, Alkharusi et al. (2014) conducted research with 411 secondary school students in Oman. They concluded that students who perceived the assessment environment as learning-oriented also reported a strong alignment with classroom instruction and high levels of authenticity, student consultation, and diversity in the assessment process.

In the Moroccan context, El Hadari and Khartite (2024) examined the perspectives of both professors and students regarding assessment practices in higher education, with a specific focus on English departments. Their study aimed to explore three main aspects: the extent to which end-of-term examinations support teaching and learning, the key challenges that impede effective assessment for learning, and the potential for adopting alternative assessment methods to enhance assessment quality. One of the study's major findings was a significant lack of alignment between course objectives, instructional approaches, and the assessment tools used to evaluate student learning. More specifically, while the instructional framework is modular, the assessment methods employed are predominantly summative and linear. Additionally, the study revealed that the practice of sharing feedback with students is almost absent.

In a slightly different study, Bouziane and Rouijel (2024) investigated the perceived benefits of assessment among Moroccan university students. Their findings indicated that although most students believed the tests they took met the standards of practical assessment, some

expressed doubts about their quality and relevance. The study also highlighted instances of good practices in alternative assessment and feedback. However, some students reported that they did not experience concrete benefits from their assessments. In other words, there is a gap between assessment design and its intended educational impact.

Students' perception of assessment and learning outcomes

In addition to examining assessment practices and testing, research has also focused on the relationship between students' perceptions of assessment tasks and performance. Leeuwenkamp et al. (2019) explored how students' perceptions of assessment quality were related to learning approaches and outcomes. They used six criteria to assess students' perceptions: the effects of assessment on learning, fairness, conditions, interpretation of test scores, authenticity, and credibility of assessment. The study revealed that students' perceptions of the conditions of assessment were correlated positively with their learning outcomes.

Research stresses the significant impact of students' perceptions on their academic performance. De Kleijn et al. (2012) found that a strong interpersonal closeness with supervisors positively correlates with improved learning outcomes, including higher final grades. Similarly, Brown and Hirschfeld (2008) demonstrated that students who perceive assessment as a means of taking responsibility for their learning tend to achieve better academic results. These two studies highlight that students' views of interpersonal relationships and assessment practices influence their motivation, learning strategies, and overall performance. Positive perceptions— such as seeing assessment as a constructive tool and feeling supported by supervisors—can lead to enhanced academic achievement. In the Moroccan context, existing research on assessment has primarily focused on the perspectives of students and teachers. However, to the researcher's knowledge, no studies have explored the relationship between students' perceptions of assessment and their performance in end-of-term examinations within the Departments of English Studies at Moroccan Faculties of Letters and Humanities. The present study was undertaken to address this gap in the literature by examining whether students' perceptions of assessment have a measurable impact on their academic achievement.

Materials and Methods

Participants

The participants in the present study were Moroccan university students from the departments of English studies in the faculties of letters and human sciences at Moulay Ismail University in Meknes, Hassan II University Ain Chock in Casablanca, and Ibn Tofail University in Kenitra. Five hundred and twenty-six students from these institutions filled out the Students' Perception of Assessment Questionnaire (SPAQ) during the fall term exams of the academic year 2022/2023. Purposive sampling was selected as the most appropriate method for this study, as it allowed for the intentional selection of students with relevant experiences and insights into assessment practices in Moroccan higher education. Specifically, students filled out the questionnaire immediately after finishing their exams to ensure that their responses reflected their

perceptions of the assessments conducted during the same academic year. This approach helped ensure that the data collected was contextually grounded and closely aligned with the research objectives.

Research Instruments

The researchers employed the Students' Perception of Assessment Questionnaire (SPAQ) as the primary instrument in this study to investigate students' perceptions of the assessment process across six key dimensions: validity, reliability, authenticity, practicality, and two aspects of the washback effect—namely, (1) the effect of assessment on study methods and (2) the effect of assessment on the content studied. The questionnaire, originally developed by Dorman and Knightley (2006), was adapted for the purposes of this study and comprised 28 items.

Before the final administration of the questionnaire, a pilot study was conducted with a group of students selected from the study sample. The questionnaire was administered to 15 students from Hassan II University of Ain Chock. Based on the feedback obtained during the pilot phase, some modifications were made to the students' questionnaire. These changes primarily involved rewording certain items and adjusting the scale format. Specifically, the word 'assessment' was replaced with 'tests' or 'testing', and the original response scales were completely removed.

To enhance the validity and reliability of the instrument used in the study, Cronbach's Alpha was calculated. The result was highly satisfactory, with a value of 0.84 for the Students' Perception of Assessment Questionnaire. This value indicates a relatively high level of internal consistency, suggesting that the instrument is reliable and appropriate for data collection.

Findings

The data collected in this study were analyzed using descriptive statistics to show students' overall perceptions as measured by the questionnaire. The five scales were validity, reliability, authenticity, practicality, and washback effect. In addition, multiple regression of these scales' scores on students' grades was calculated to measure the relationship between students' perception and performance.

Students' perception of assessment tasks

Table one presents students' perceptions of the validity of assessment tasks. The first five items on the SPAQ questionnaire in Table one addressed the validity of assessment tasks in the departments of English studies at the three faculties under investigation. Items one, two, and three evaluated the congruence of assessment tasks with planned learning or content validity. In addition, items four and five examined the face validity of the tests administered to students in these departments. The scores of participants' responses in the three faculties were almost similar and exceeded (M = 3.35) for content validity and (M = 3.50) for face validity. This result indicates that students perceived the assessment tasks designed by their professors as similar to what they do in class. Students also perceived the tests designed by their professors in the three faculties as having a satisfactory level of face validity.

Table 1. *Students' perception of the validity of assessment tasks*

	Moulay Ismail Faculty			Ain Chock Faculty			Ibn Tofail Faculty		
	Num	Mean	SD	Num	Mean	SD	Num	Mean	SD
1. Exams at my Faculty are about what I have done in class.	207	3.63	1.02	159	3.45	1.02	160	3.63	0.85
2. I am tested on what the teachers have taught me during the semester.	207	3.90	0.91	159	3.68	1.03	160	3.93	0.91
3. Exam questions are about topics that have been covered in class.	207	3.61	0.97	159	3.35	1.00	60	3.63	0.91
4. The formats of the tests are expected and well constructed with familiar tasks.	207	3.61	1.06	159	3.52	1.05	160	3.92	0.82
5. The items in the tests are straightforward, uncomplicated, and can be accomplished within the allotted time limit.	207	3.61	1.06	159	3.52	1.05	160	3.78	0.96

Note. SD: Standard Deviation

Table two presents students' perceptions of the reliability of assessment tasks at the three faculties. According to the scores in this table, all the mean values exceeded (M = 3.10), which means that students considered the assessment tasks at the three Faculties reliable.

Table 2. *Students' perception of the reliability of assessment tasks*

	Moulay Ismail Faculty			Ain Chock Faculty			Ibn Tofail Faculty		
	Num	Mean	SD	Num	Mean	SD	Num	Mean	SD
1. The tests I sit for have clear directions for scoring/ evaluation.	207	3.35	1.16	159	3.20	1.14	160	3.45	1.19
2. They have uniform rubrics for scoring/ evaluation.	207	3.60	1.13	159	3.31	1.13	160	3.61	1.07
3. They contain items/ tasks that are unambiguous to me.	207	3.36	1.03	159	2.99	1.07	160	3.10	1.16

Table three presents students' perceptions of the authenticity of assessment tasks. Among the 5 Likert scales in this table, all authenticity items exceeded half of it (M = 3). Accordingly, it could be claimed that the assessment tasks at the three faculties were perceived as authentic.

Table 3. *Students' perception of the authenticity of assessment tasks*

	Moulay Ismail Faculty			Ain Chock Faculty			Ibn Tofail Faculty		
	Num	Mean	SD	Num	Mean	SD	Num	Mean	SD
1. I am asked to apply my learning to real-life situations.	207	3.55	1.09	159	3.45	1.14	160	3.71	1.03
2. My English final exams are meaningful and useful.	207	3.36	1.12	159	3.10	1.15	160	3.16	1.19
3. I find exam tasks relevant to the real world.	207	3.56	1.11	159	3.21	1.04	60	3.54	1.07
4. English final exams test my ability to apply learning.	207	3.53	1.10	159	3.51	1.00	160	3.70	1.02

5. Exams at university examine my ability to answer every day questions.	207	3.16	1.09	159	3.23	0.99	160	3.38	1.03

The results in Table four present students' perceptions of the practicality of assessment tasks. Unlike the previous criteria, which generated similar results among participants in the three faculties of letters and humanities under study, the item's mean score at Ain Chock Faculty indicated that the students perceived the assessment tasks as impractical. Items three and four, for instance, scored less than (M = 2.50). This result indicates that, according to students, exams at this Faculty were not conducted smoothly and lacked sufficient resources to ensure proper testing conditions.

Table 4. *Students' perception of the practicality of assessment tasks*

	Moulay Ismail Faculty			Ain Chock Faculty			Ibn Tofail Faculty		
	Num	Mean	SD	Num	Mean	SD	Num	Mean	SD
1. All the printed materials are accounted for.	207	2.96	1.19	159	2.85	1.27	160	3.43	1.16
2. The allotted time is enough to finish all the exam questions.	207	2.83	1.08	159	2.97	0.99	160	3.25	1.05
3. The administration of exams is carried out smoothly and efficiently.	207	3.14	1.00	159	2.48	1.02	60	3.19	1.07
4. There are enough teachers to invigilate in exams.	207	3.59	0.97	159	2.43	1.00	160	3.55	1.12
5. There are enough materials and resources to administer the tests in good conditions.	207	3.28	1.20	159	3.08	1.20	160	3.65	0.85

Table five presents the scores related to students' perceptions of the effect of assessment tasks on their studied content at the three Faculties under study. The results show that the mean of score items exceeded (M = 2.50), with the lowest mean score being (M = 3.43). This result indicates that students at the three institutions agree that the assessment tasks designed by their professors affect the content they study.

Table 5. *Students' perception of the effect of assessment on their studied content*

	Moulay Ismail Faculty			Ain Chock Faculty			Ibn Tofail Faculty		
	Num	Mean	SD	Num	Mean	SD	Num	Mean	SD
1. All sections of courses/ modules are of equal importance to me.	207	3.61	1.00	159	3.51	1.21	160	3.96	0.89
2. When studying my courses, I focus more on the final exam than learning goals.	207	3.89	0.84	159	3.30	1.16	160	3.63	0.93
3. I prioritize learning the reviewed points in courses based on the priorities of final exams.	207	3.71	0.95	159	3.43	0.99	60	3.93	0.84

4. The supplementary materials that I study are influenced by the final exams.	207	3.63	1.01	159	3.51	1.04	160	3.78	0.97
5. I expect my professors to place greater emphasis on the topics that will be covered in the final exams.	207	3.48	0.92	159	3.51	1.07	160	3.47	1.18

The last Table in this section (Table six) presents the scores related to student' perceptions of the effect of assessment tasks on their studying methods at the three faculties of letters and humanities. The results show that all the means of score items exceeded (M = 2.50), with the lowest score being (M = 3.13). This result indicates that students at the three faculties agree on the effect of assessment tasks on the studying method they adopted.

Table 6. *Students' perception of the effect of assessment on their studying method*

	Moulay Ismail Faculty			Ain Chock Faculty			Ibn Tofail Faculty		
	Num	Mean	SD	Num	Mean	SD	Num	Mean	SD
1. Final examinations influence my method of studying.	207	3.33	1.06	159	3.62	0.98	160	3.59	1.08
2. I try to study based on previous final exams.	207	3.46	1.10	159	3.32	1.11	160	3.13	1.17
3. I spend less time on sections that are less likely to appear in the final exams.	207	3.58	0.89	159	3.71	1.12	60	3.75	1.03
4. I study based on the final examinations.	207	3.45	1.10	159	3.45	1.11	160	3.17	1.20
5. I try to learn test-taking strategies.	207	3.49	1.22	159	3.51	1.15	160	3.51	1.09

The relation between students' perception of assessment and learning outcomes (grades)

Parameter estimates: Age, Faculty, and Gender

This section deals with parameter estimates of age, Faculty, and gender on language assessment principles (authenticity, practicality, validity, reliability, and washback effect).

Validity

Since validity has a very significant effect on the quality of the tests, it was also importance to study the effect of Faculty, gender, and age on validity items. As displayed in Table seven, the impact of Faculty on these items was highly significant for items two and three at Ain Chock Faculty and item 4 at Ibn Tofail Faculty. Gender effects on validity items two and five was also substantial; however, there was no effect of age on validity items.

Table 7. *Faculty, gender and, age effects on validity items and their Nagelkerke index*

	Item1		Item2		Item3		Item4		Item5	
	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]
Ain Chock	-0.43	0.28	-0.45	0.01	-0.53	0.006	-0.22	0.24	-0.36	0.06
Ibn Tofail	-0.13	0.49	0.09	0.63	0.05	0.77	0.51	0.01	0.13	0.49

ISSN: 2229-9327

Gender	0.26	0.10	0.34	0.03	0.09	0.55	0.27	0.09	0.42	0.008
Age	-0.02	0.36	0.02	0.34	0.004	0.90	0.02	0.34	-0.05	0.06
Nagelkerke Index	0.01		0.02		0.02		0.03		0.03	

Note. [a]Est: Estimte, [b]p: P value. Bold values are significant at a 5% probability.

Reliability

Another primary criterion of language assessment investigated in this study was reliability. Three items were included in the questionnaire designed for students in the three Faculties under study. Table eight represents faculty, gender, and age effects on items' reliability and Nagelkerke index. The results revealed a significant impact on item two reliability at Ain Chock Faculty and item three reliability at Ibn Tofail Faculty. However, gender and age did not affect reliability items.

Table 8. *Faculty, gender, and age effects on reliability items and their Nagelkerke index*

	Item1		Item2		Item3	
	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]
Faculty Ain Chock	-0.63	0.001	-0.09	0.63	0.14	0.45
Faculty Ibn Tofail	-0.13	0.48	0.28	0.14	0.38	0.04
Gender	0.10	0.50	-0.06	0.67	-0.11	0.46
Age	-0.04	0.15	-0.04	0.09	-0.03	0.18
Nagelkerke index	0.02		0.01		0.01	

Authenticity

Table 9 presents the effects of Faculty, gender, and age on authenticity item' and Nagelkerke index. The results showed that the p-values of items 3 and 5 at Ain chock Faculty were less than 0.05, which means that the effect of this Faculty on these two items' authenticity was significant. However, the p-value of all the other items at Ibn Tofail Faculty was more than 0.05, which means that this Faculty had no effect on any items' authenticity and their Nagelkerke index. Finally, the impact of age on item 4 authenticity was also significant (p = 0.04).

Table 9. *Faculty, gender, and age effects on items' authenticity and their Nagelkerke index*

	Item1		Item2		Item3		Item4		Item5	
	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]
Ain Chock	-0.30	0.10	-0.52	0.06	-0.57	0.003	-0.20	0.28	-0.40	0.03
Ibn Tofail	0.10	0.59	-0.10	0.58	-0.32	0.09	0.26	0.17	-0.27	0.15
Gender	-0.30	0.60	0.54	0.73	0.07	0.62	0.13	0.29	0.11	0.46
Age	-0.01	0.54	-0.20	0.49	0.03	0.22	-0.06	0.04	0.02	0.45
Nagelkerke Index	0.016		0.17		0.23		0.20		0.12	

Practicality

Table 10 presents Faculty, gender, and age effects on practicality items and their Negelkerke index. The results showed that the impact of Ain Chock Faculty was highly significant on items

ISSN: 2229-9327

three and four and significant on item five practicality. In addition, the impact of Ibn Tofail on the practicality of Items one and two was highly significant. As for gender, its effect on item four was significant, whereas gender did not affect other practicality items. Finally, the impact of age on practicality items was highly significant for items two and five.

Table 10. *Faculty, gender, and age effects on practicality items and their Nagelkerke index*

	Item1		Item2		Item3		Item4		Item5	
	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]
Ain Chock	-0.16	0.37	0.25	0.17	-1.26	0.00	-2.07	0.00	-0.37	0.05
Ibn Tofail	0.71	0.00	0.70	0.00	0.04	0.83	-0.18	0.37	0.50	0.01
Gender	-0.06	0.67	0.09	0.95	0.10	0.51	0.32	0.05	0.27	0.09
Age	-0.02	0.39	-0.06	0.02	0.01	0.63	-0.01	0.66	-0.07	0.01
Nagelkerke Index	0.45		0.04		0.10		0.21		0.05	

The Washback

The last assessment principle investigated in the present study was the washback effect. In this study, this criterion of language assessment was dealt with in two different ways: the washback effect on students' studied content and students' studying method.

The Washback (effect on studied content)

The results in Table 11 showed that the p-value of items 2 and 3 at Ain Chock Faculty and items 2 and 5 at Ibn Tofail Faculty was less than 0.05, meaning a highly significant effect of the two faculties on the cited washback items. The results also revealed that age had a considerable effect on washback item four.

Table 11. *Faculty, gender, and age effects on items' washback (content) and their Nagelkerke index*

	Item1		Item2		Item3		Item4		Item5	
	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]
Ain Chock	-0.05	0.78	-1.05	0.00	-0.64	0.001	-0.27	0.15	0.14	0.45
Ibn Tofail	0.004	0.98	-1.29	0.00	-0.32	0.10	-0.33	0.08	0.45	0.02
Gender	-0.24	0.12	0.08	0.60	0.02	0.86	-0.03	0.85	0.10	0.53
Age	0.03	0.24	0.01	0.64	-0.01	0.57	-0.05	0.05	0.001	0.97
Nagelkerke Index	0.008		0.10		0.02		0.012		0.013	

The Washback (effect on studying method)

The last table in this section concerned Faculty, gender, and age effects on washback items (effect on studying method). The results displayed in this table (Table 11) showed a highly significant effect of Ain Chock Faculty on washback item one as well as a substantial effect of Ibn Tofail Faculty on the washback items one, two, and three. Similarly, gender had a substantial effect on washback item two. Finally, Age had a significant effect on washback item four.

Table 12. *Faculty, gender, and age effects on items' washback (method) and their Nagelkerke index*

	Item1		Item2		Item3		Item4		Item5	
	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]	Est[a].	P[b]
Ain Chock	0.54	0.06	-0.19	0.31	0.36	0.06	-0.03	0.87	-0.01	0.95
Ibn Tofail	0.84	0.00	0.45	0.02	-0.44	0.02	0.32	0.10	0.20	0.29
Gender	0.18	0.26	-0.39	0.01	0.11	0.46	-0.11	0.48	0.12	0.44
Age	0.05	0.07	0.009	0.76	0.03	0.24	-0.07	0.008	-0.03	0.21
Nagelkerke Index	0.04		0.032		0.033		0.025		0.009	

Instead of dealing with the 28 items, standardization of the variables allowed us to deal with only five concepts. This data reduction was conducted using cluster analysis under the distance ward method, as demonstrated in the Dendrogram below. The discrimination thread in this Dendrogram was around seven, which resulted in well-defined clusters based on which we used the scores displayed in Table 12.

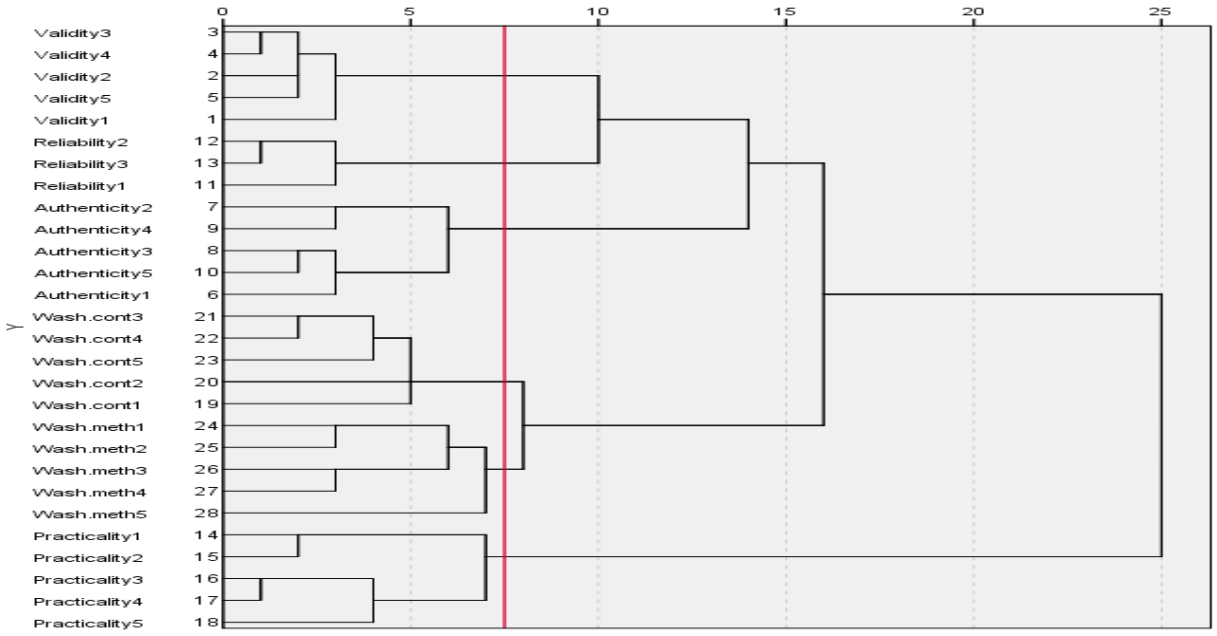


*Figure 1***.** Dendrogram representing Ward distance

Table 13. *Effect of faculty, gender, and age on validity, authenticity, practicality, reliability, and washback scores*

	Validity		Authenticity		Practicality		Reliability		Washback	
Source	Obs. F	P value	Obs. F	P value	Obs. F	P value	Obs. F	P value	Obs. F	P Value
Faculty	8,615	0,001	5,318	0,005	28,733	0,001	2,130	0,120	1,774	0,171
Gender	8,055	0,005	2,444	0,119	1,421	0,234	0,009	0,923	0,025	0,874
Age	0,166	0,684	0,166	0,683	3,507	0,062	5,508	0,019	0,010	0,919

The results in Table 13 showed a highly significant effect of Faculty on validity, authenticity, and practicality scores. In addition, gender also had a significant impact on validity scores. However, age did not affect any language assessment criteria, which were also used in this study as an element of students' perception of assessment.

The last table in this section (Table 14) presents multiple regressions of validity, authenticity, practicality, reliability, and washback scores on students' grades. The results showed that apart from validity, which had a significant effect on students' grades, the other elements representing students' perception of assessment did not impact grades.

Table 14. *Multiple regression of validity, authenticity, practicality, reliability, and washback scores on students' grades*

	Unstandardized Coefficients		Standardized Coefficients		
Model	B	Std. Error	Beta	T	P value
Validity	0,148	0,050	0,141	2,955	0,003
Authenticity	-0,057	0,044	-0,063	-1,275	0,203
Practicality	0,065	0,041	0,075	1,587	0,113
Reliability	0,051	0,069	0,036	0,738	0,461
Washback score	-0,033	0,035	-0,043	-0,938	0,348

To explain the nature of the relationship between validity and students' grades, for instance, the following graph represents a simple scatter of grades by validity.

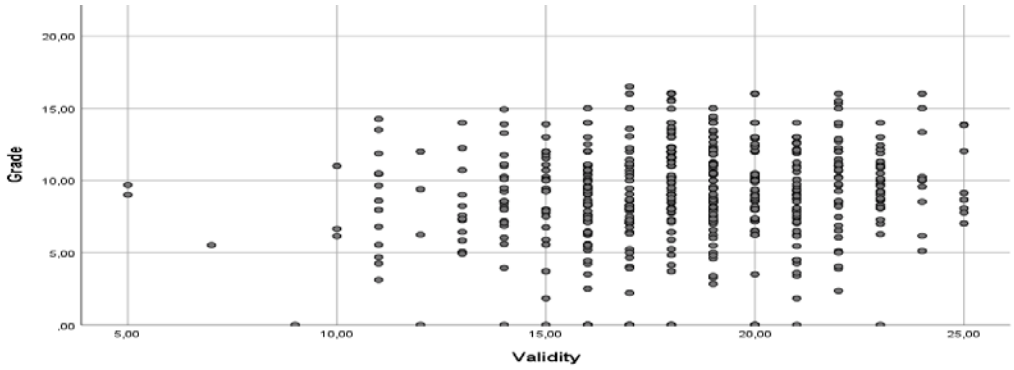


Figure 2. Simple Scatter of Grade by Validity

The distribution of the points on the scatter, which are dispersed and do not follow any particular direction, provides a visual representation of the relationship between the validity of assessment tasks and students' grades.

Discussion

This study addresses two critical questions related to how Moroccan university students perceive the assessment tasks designed by their professors and how this perception relates to their academic performance. As previously mentioned, validity, reliability, authenticity, practicality, and washback were the five key factors influencing students' perceptions of assessment. The

Washback concept was examined regarding how assessments impact students' study methods and the content they focus on.

In response to the first research question, the average scores of students on the five-scale assessment questionnaire indicate that none of the items received the maximum score. However, most scales reached acceptable scores in the three faculties. The study, for instance, showed that students perceived authenticity to have the highest mean in the three faculties (M = 3.67 at Moulay Ismail faculty, M = 3.50 at Ain Chock Faculty, and M = 3.77 at Ibn Tofail Faculty). In fact, "there has long been an agreement over the idea that authenticity is an important quality for the test development." (Lynch, 1982, p. 11). Morrow (1991) emphasizes the paramount importance of authenticity, while Wood (1993) also regards it as one of the most crucial aspects of language testing. Bachman and Palmer (1996, p. 23) describe authenticity as a "critical quality of language tests." Gulikers et al. (2008) propose that authenticity is subjective and often tied to an individual's personal experience. Students' perceptions of the authenticity of assessment tasks align with the findings of Gao (2012), who argues that for assessments to reflect real-life situations, it is essential that teachers and other stakeholders involved in assessment design fully understand the real-world contexts that are most relevant to students. In contrast, studies by Dhindsa (2007) and Rahman (2020) reported a perceived lack of authenticity in assessment tasks, as reflected in their investigation of students' perception of assessment tasks.

Alongside 'Authenticity,' students in the English studies departments at the three faculties also rated 'Validity' relatively high (M = 3.67 at Moulay Ismail Faculty, M = 3.50 at Ain Chock Faculty, and M = 3.77 at Ibn Tofail Faculty) compared to other scales. This result suggests a strong alignment between assessment tasks and classroom instruction. Many participants indicated that the assessments they undertake cover topics that were discussed in class. When students know they will be tested on the material they have studied, they tend to feel more motivated and eager to invest time and effort in class-related activities (Brookhart & Bronowicz, 2003). Similarly, in a study on students' perceptions of Mathematics tasks, Gao (2012) highlighted the importance of informing students about how assessments align with the goals, objectives, and activities of their learning program. However, when assessments are not aligned with these goals and the teaching process, it can undermine students' learning (Segers, 2009).

Practicality had the lowest scores, especially at Ain Chock Faculty (M = 2.76). In this respect, Bachman and Palmer (1996) caution that practicality should be given the same importance as the other assessment criteria when designing language tests. The complexity of investigating the practicality of assessment tasks is related to the number of stakeholders associated with this criterion. The practicality of language tests is reflected in the perspectives of test takers, test designers, and the conditions under which the tests are administered. According to Brown and Abeywickrama (2010), a practical test considers the time and effort required for both its design and scoring. Additionally, it should be manageable for the test taker, allowing completion within appropriate time limits.

Practicality was examined within the broader framework of Test Usefulness, which comprises six key criteria for language assessment. Bachman and Palmer (1996), who developed this framework, defined practicality as the balance between the resources needed for designing, developing, and administering the test and the resources available for these activities. They identified three key types of resources necessary to ensure the practicality of language tests: human

resources (such as test administrators, writers, and scorers), material resources (including word processors, tape and video recorders, and computers), and time (for both test development and completion of specific tasks). However, to our knowledge, no studies have been published on the practicality of assessment tasks within the Moroccan context.

Reliability is another crucial aspect of assessment quality. A test can only be considered valid if it is reliable (Al Noor, 2020). Despite the different methods used to measure the reliability of assessment tasks, researchers agree that reliability refers to the consistency of test scores (Harris, 1969; Gronlund, 1977; Coombe et al., 2007; Brigui, 2017). Brown and Abeywickrama (2010) determined three criteria for a test to be reliable. The test should have clear directions, uniform rubrics for scoring, and clear items and tasks. Accordingly, the results in Tables 4, 5, and 6 showed that students in the three faculties perceived the assessment tasks as reliable. The scores were (M=3.43) at Moulay Ismail Faculty, (M=3.16) at Ain Chock Faculty, and (M=3.49) at Ibn Tofail Faculty. This fact means that the assessment tasks at these faculties have clear directions and uniform rubrics for scoring. In addition, these tasks contain relatively clear items. Some factors may affect the reliability of language tests. Sattler (2001) emphasized that test length is a key factor in determining the reliability of tests, asserting that longer tests tend to be more reliable. Hughes (2003) identified scoring as another source of unreliability in language tests. Additionally, Li (2019) emphasized that test conditions can significantly influence test outcomes.

The last criterion investigated in this study was the washback effect on students' study content and method. Regarding the first type of washback, results showed that the assessment tasks affect students' study methods. In the three faculties, the scores of the five items included in this scale exceeded (M = 3.00). This result means that students in the three faculties were more concerned with getting good grades in final exams than with pedagogical aims. For this reason, they focused more on points that were more likely to appear in final exams. The items used to measure the effect of assessment tasks on studying content generated similar results in this study. All the items in the three faculties also exceeded (M = 3.30), indicating that the assessment tasks affect students' studying content. Students often approach their studies and attend classes with previous final exams in mind, ultimately dedicating more time and effort to topics they believe are most likely to appear on these exams.

These findings align with the results of many other studies (e.g., Andrews et al., 2002; Alderson & Wall, 1993; Cheng & Curtis, 2004; Ferman, 2004; McNamara, 2000; Hemmati & Soltanpoor, 2014), all of which concluded that the washback effect could significantly influence learners. These findings supports Messick's conception of validity, which includes both the value promoted by the test and its impact on individuals (Fulcher, 2010). In other words, students tailored their study methods around previous final exams, focusing more on sections likely to appear in future tests. As a result, their study strategies would have changed entirely if the design and structure of the final exams had been different.

The findings in this study were supported by other studies investigating the washback effect. Alderson and Hamp-Lyons (1996) conducted a study that examined the influence of the Test of English as a Foreign Language (TOEFL) on classroom teaching. The researchers concluded that TOEFL affected the content of the course. Hemmati and Soltanpour (2014), whose study aimed to investigate the washback effect of university-level final examinations on Iranian

distance learners' study methods and content, reached similar results. These results indicated that final exams had a significant impact on students, influencing both the content they studied and the methods they used.

Our results indicate that students across the three faculties perceive all sections of their courses or modules as equally important (M = 3.61 at Moulay Ismail Faculty, M = 3.51 at Ain Chock Faculty, and M = 3.96 at Ibn Tofail Faculty). This perception highlights the inherent significance of all course content. Students' overall perceptions are most likely influenced by the nature of the final exams. From the researchers' experience as educators, it is common practice to inform students that any part of the course may be included in the exams, contributing to their perception of uniform importance. In other institutions and due to the methods of assessment design, teachers find themselves testing details and marginal points. In Ibn Tofail University, for instance, exams designed for common core students use only multiple-choice questions (Brigui, 2017). Because some tests are designed using more than 30 questions, teachers may find themselves testing most of the points included in the course sections. These practices confirms what Alderson & Wall (1993) stated regarding the influence of testing on teaching and learning, noting that a test causes language teachers and learners to do things "they would not necessarily otherwise do because of the test." (p. 117).

Concerning the second research question, the study found that students' overall perception of assessment negatively affected their learning outcomes. The only factor influencing assessment scores was validity, though it explained only a small portion of the variance in students' performance. The remaining factors—authenticity, reliability, practicality, and washback—did not significantly correlate with students' grades. This lack of correlation might be because students do not prioritize assessment criteria, leading to a disconnect between their perceptions and actual behaviors. Leeuwenkamp and Brink (2007) argue that not all assessment criteria are crucial for ensuring student satisfaction. Additionally, students may not fully understand these criteria or how they impact their perceptions and performance. To gain a deeper understanding of how students form perceptions of assessment quality, further research is needed to explore these views in greater depth, which aligns with the findings of Bouziane and Rouijel (2024), whose study revealed that although students generally regard their assessments as satisfactory, several key areas—such as validity, feedback, and the overall impact on learning—still require significant improvement.

The methodology adopted in this study may have influenced the results obtained. Firstly, students were asked to provide their perceptions of assessment tasks in general without specifying particular programs or modules. Future research could investigate whether students' perceptions of assessment quality vary across different modules. Secondly, the study sample included students from various levels of study (semesters 1 through 4) and other faculties of letters and humanities at three Moroccan universities (Moulay Ismail University, Hassan II University, and Ibn Tofail University). To obtain more precise insights, future research could apply the SPAQ to specific Faculties and targeted assessment programs.

Other factors that could account for the results of this study may be specific to the Moroccan context. To better understand the relationship between students' perceptions of assessment quality and their grades in Morocco, it is essential to focus on the Moroccan educational context. Relating students' performance of assessment tasks to the assessment quality criteria might not give us a clear relationship between their perception and performance. Students may sometimes have a negative perception of assessment quality but still get high grades, and vice versa. Additionally, participants' attitudes toward scientific research can also play an important role and potentially influence the results. Furthermore, students might not have read the items in

the questionnaires attentively. More attention should, therefore, be given to the choice of participants and piloting of the questionnaire to ensure that students know the study items well.

Conclusion

This study aimed to gain deeper insights into Moroccan university students' perceptions of assessment tasks in the English studies departments at Moroccan faculties of letters and humanities. It also examined the relationship between students' perceptions of assessment and performance. The findings revealed that students perceived the assessment tasks at the three faculties as having a satisfactory level of validity, reliability, and authenticity. However, the practicality of these tasks was questioned, with notable dissatisfaction reported by students, particularly at the Ain Chock Faculty. Additionally, the study found that the assessment tasks designed by Moroccan professors influenced both the content students studied and their study methods.

Surprisingly, there was a discrepancy between students' overall perceptions of assessment tasks and testing across the three faculties and their learning outcomes. This finding indicates that not all variables related to students' perceptions of assessment tasks are related to their performance or learning outcomes. Moreover, the variables that do relate to students' learning outcomes account for only a small portion of the variance. While these results do not diminish the importance of students' perceptions of these variables, they highlight the need to integrate students' views on all assessment criteria to ensure the quality of testing and assessment practices in Moroccan higher education.

Funding: This research in not funded.
Acknowledgments: Not applicable.
Conflicts of Interest: The authors declare no conflict of interest.
Authenticity: This manuscript is an original work.
Artificial Intelligence Statement: AI and AI-assisted technologies were not used.

About the Authors

Dr. Anouar Mohamed El Kasri, PhD in Applied Linguistics from Moulay Ismail University in Meknes, Morocco, has been a teacher of English for over 14 years. He has actively participated in many national and international conferences. His primary research interests include Teaching English as a Foreign Language (TEFL), assessment, and soft skills.
ORCID ID: https://orcid.org/0009-0001-9485-9777

Dr. Moulay Sadik Maliki has been a professor of English linguistics at the faculty of arts, and humanities, Hassan II University, Casablanca, since 1988. He has published many articles and books on intercultural communication, culture and linguistic diversity in Morocco. Dr Maliki is interested in cultural issues, applied linguistics, and translation.
ORCID ID: https://orcid.org/0009-0008-4246-4395

Dr. Mohammed Larouz is the Dean of the Faculty of Letters and Humanities at Moulay Ismail University in Meknes, Morocco, where he has been a faculty member since 2005. His research areas and expertise include applied linguistics issues in Morocco, and he has research interests in TEFL, sociolinguistics, research methodology, and communication.

Dr. Martina Pavlikova is a university teacher, senior lecturer at the Department of Journalism, Constantine the Philosopher University in Nitra, Slovakia, where She completed her doctoral studies in Cultural Studies. Her research is devoted to the literature of English-speaking countries, ethical behaviour in the mass media and modern trends in visual arts.
ORCID ID: https://orcid.org/0000-0002-6738-3320

Dr. Brahim El Yousfi is a former Scientist at Plant Pathology Lab, National Institute of Agricultural Research (INRA), Morocco. He got his degree of engineer in phytopathology from Minnesota University, USA in 1995. Dr EL Yousfi published more than 30 peer-reviewed papers and was an expert for several projects and organizations.
ORCID ID: https://orcid.org/0000-0003-1806-8719

References

Ajmal, M., Maqbool, S., Pavlikova, M., Alam, S., & Banu, S. (2024). Investigating Teacher-Student Emotional Dynamics in English Language Teaching: Examining Co-Regulation Techniques and Their Impact on Learning Outcomes. *World Journal of English Language, 15*(3), p215. doi:http://dx.doi.org/10.5430/wjel.v15n3p215

Alderson, J. C., & Wall, D. (1993). Does washback exist? Applied Linguistics, *14*(2), 115-129. http://dx.doi.org/10.1093/applin/14.2.115

Alderson, J. C., & Hamp-Lyons, L. (1996). TOEFL preparation courses: A study of washback. Language Testing, *13*(3), 280–297.

Alkharusi, H. (2013). Canonical correlational models of students' perceptions of assessment tasks, motivational orientations, and learning strategies. *International Journal of Instruction*, *6*, 21-38.

Alkharusi, H., Aldhafri, S., Alnabhani, H., & Alkalbani, M. (2014). *Classroom Assessment: Teacher Practices, Student Perceptions, and Academic Self-Efficacy Beliefs. Social Behavior and Personality*, *42*(5), 835-856.

Al Noor, H. (2020). A Probe into the different aspects of 'Validity' and 'Reliability' of the IELTS writing test. *International Journal of English Literature and Social Sciences*, *5*(4)

Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford.

Bachman, L. F., & Palmer, A. S. (1996). Language testing in practice: Designing and Developing seful language tests. Oxford: Oxford University Press.

Beg, B., Amir, A., Alam, S., Kralik, R., & Warda, W. (2024). Investigating the Efficiency of the Rotation Model in Improving First-Year Undergraduate ESL Learners' Writing: A Quasi-Experimental Study. *World Journal of English Language, 15*(2), 357-367. doi:http://dx.doi.org/10.5430/wjel.v15n2p357

Black, P., & William, D. (1998). *Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice*, *5*, 7-74.

Brigui, H. (2017). Investigating the ITU semi-computerized MCQ model: A systematic assessment of EFL students' attitudes and testing behaviors in the newly adopted e-

testing form. IOSR Journal of Research & Method in Education (IOSR-JRME). *7*(3), 19-25.

Bouziane, A., Rouijel, E. (2024*).* Assessment for Learning in Higher Education in Morocco: The Students' Perceptions. *European Journal of Education Studies, 11*(11), 428-443. http://dx.doi.org/10.46827/ejes.v11i11.5615

Brookhart, S. M., & Bronowicz, D. L. (2003). *"I don't like writing. It makes my fingers hurt": Students talk about their classroom assessments*. Assessment in Education, *10*(2), 221–241.

Brown, G. T. L., & Hirschfeld, G. H. F. (2008). *Students' conceptions of assessment: Links to outcomes. Assessment in Education: Principles, Policy & Practice*, *15*, 3–17.

Brown, H. D., & Abeywickrama, P. (2010). *Language Assessment: Principles and Classroom Practices (3rd ed.)*. London: Pearson.

Brown, H. D., & Abeywickrama, P. (2010). *Language Assessment: Principles and Classroom Practices (3rd ed.)*. London: Pearson.

Brown, S., Rust, C. & Gibbs, G. (1994). *Strategies for diversifying assessment in higher education*. Oxford: The Oxford Centre for Staff and Learning Development.

De Kleijn et al., (2012). *Relationship, final grade, perceived supervisor contribution to learning and student satisfaction Master's thesis supervision: Relations between perceptions of the supervisor–Student*. Studies in Higher Education, *37*, 925–939.

Dhindsa, H., Omar, K., & Waldrip, B. (2007). Upper secondary Bruneian science students' perceptions of assessment. *International Journal of Science Education, 29*, 1261- 1280.

Dochy, F.; McDowell, L. (1997). *Assessment as a tool for learning. Studies in Educational Evaluation*, *25*(4).

Dorman, J. P., & Knightley, W. M. (2006). Development and validation of an instrument to assess secondary school students' perceptions of assessment tasks. Educational Studies, *32*(1), 47–58. https://doi.org/10.1080/03055690500415951

Coombe, C., Folse, K., & Hubley, N. (2007). A practical guide to assessing English language learners. Ann Arbor: The University of Michigan Press.

Ekbatani, G. V. (2000). *Moving Toward Learner-Directed Assessment*. In Ekbatani, G.V., & Pierson, H. V. (eds.), *Learner-directed Assessment in ESL* (pp. 1-13). Laurence Erlbaum: Mahwah NJ

El Hadari, E., & Khartite, B. (2024). *Professors and Students' Perspectives on Assessment in Moroccan Higher Education: The Case of English Departments*. International Journal of Interdisciplinary Gender Studies. *5*(1), 1-17

Fulcher, G. (2010). *Practical language testing*. London: Hodder Education, & Hachette UK company.

Gibbs, G. (1999). *Using assessment strategically to change the way students learn*. In Brown, S., & Glasner, A. (Eds.), *Assessment matters in higher education: Choosing and using diverse approaches* (pp. 41-53). Buckingham

Gao, M. (2012). *Classroom Assessment in Mathematics: High School Students Perceptions*. International Journal of Business and Social Science, *3*(2), 63-68

Gielen, S., Dochy, F., & Dierick, S. (2003). *Evaluating the consequential validity of new modes of assessment: The influence of assessment on learning, including pre-, post-, and true assessment effects*. In Segers, M. F., & Cascallar, E. (Eds.), *Optimizing new modes of assessment: In search of qualities and standards* (pp. 37-54). Dordrecht: Kluwer Academic Publishers.

Gronlund, N. E. (1977). *Constructing Achievement Tests (2nd ed.)*. Englewood Cliffs: Prentice Hall.

Gulikers, J. T., Bastianens, T, J., Kirschner, P. P., & Kerster, L. (2008). Authenticity is in the eye of the beholder: Student and teacher perceptions of assessment authenticity. *Journal of Vocational Education and Training, 60* (4), 401-412.

Harlen, W., & Crick, R. D. (2003). *Testing and Motivation for Learning, Assessment in Education: Principles, Policy and Practice*, *10* (2)

Harris, M., & McCann, P. (1994). *Assessment*. Oxford: Heinemann Publishers.

Hemmati, F., & Soltanpour, F. (2014). Final Examinations and Their Washback Effects: A Study on Iranian Payame Noor University Distance Learners. *Malaysian Journal of Distance Education*, *16*(1), 15-36.

Horwitz, E. K. (1989). Facing the blackboard: student perceptions of language learning and the language classroom. ADFL Bulletin, *20*(3), 61–64.

Hughes, A. (1993). Testing for language teachers. Cambridge: Cambridge University Press.

Erickson, G., & Gustafsson, J. (2005). *Some European students and teachers' views on language testing and assessment: a report on a questionnaire survey*. Retrieved from http://www.ealta.eu.org/documents/resources/enlta%20activ ity%204%20report.pdf

Leeuwenkamp, K. J. G., Brinke, D.J., & Kester, L. (2019). Students' perceptions of assessment quality related to their learning approaches and learning outcomes. *Journal of Studies in Educational Evaluation. 63*, 72-82.

Lindblom-Ylänne, S., & Lonka, K. (2001). *Students' perceptions of assessment practices in a traditional medical curriculum*. Advances in Health Sciences Education, *6*(2), 121–140.

Linnenbrink, E. A., & Pintrich, P. R. (2001). *Multiple goals, multiple contexts: The dynamic interplay between personal goals and contextual goal stresses*. In Volet, S., & Järvelä, S. (Eds.), *Motivation in learning contexts* (pp. 251-270). Elmsford, NY: Pergamon.

Linnenbrink, E. A., & Pintrich, P. R. (2002). *Motivation as an enabler for academic success*. School Psychology Review, *31*, 313-327.

LI, J. (2019). *An Evaluation of IELTS Speaking Test*. Open Access Library Journal, *6*.

Lynch, A. J. (1982). Authenticity in language teaching: Some implications for the design of listening materials. *British Journal of Language Teaching 20*, 9-16.

Messick, S. (1996). *Validity and washback in language testing. Language Testing 13*(3). 241–

256.

Hemmati, F., & Soltanpour, F. (2014). Final Examinations and Their Washback Effects: A Study on Iranian Payame Noor University Distance Learners. *Malaysian Journal of Distance Education 16*(1), 15-36.

McMillan, J. H., & Workman, D. J. (1998). *Classroom assessment and grading practices: A review of the literature*. ERIC Document Reproduction Service No. ED453263.

Morrow, K. (1991). Evaluating communicative tests. In S. Anivan (ed.). Current developments in language testing. (pp. 111-118). Singapore: SEAMO Regional Language Center.

Mussawy, S. A. J. (2009). *Assessment practices: Students and teachers' perceptions of classroom assessment*. (Master's thesis, University of Massachusetts).

Nevo, D. (1995). *School-based evaluation: A dialogue for school improvement*. Oxford: Pergamon.

Othmane, M.. & Bouyakoub, N. (2020). Teaching and Testing English in the Algerian Educational System. *Arab World English Journal*, *11*(1) 444 – 458. DOI: https://dx.doi.org/10.24093/awej/vol11no1.30

Rahman, F. (2020). *Undergraduate Students' Perception Towards Grammar Assessment in the EFL Classroom*. Saga, 1(2).

Richards, J. C., & Rogers, T. (2016). Approaches and methods in language teaching New York: Cambridge University Press.

Sattler. J. M. (2001). *Assessment of Children Cognitive Applications (4th ed.)*. USA: Publisher Inc.

Segers, M., Dochy, F., Gijbels, D., & Struyven, K. (2009). *Changing insights in the domain of assessment in higher education: Novel assessments and their pre-, post- and pure effects on student learning*. In McInerney, D. M., Brown, G. T. L., & Liem G. A. D., (Eds.). *Student perspectives on assessment: What students can tell us about assessment for learning*, (pp. 297–319). Charlotte, NC: Information Age Publishing, INC.

Shepard, L. A. (2000). *The role of assessment in learning culture*. Education Research, *29*(7),4-14.

Van de Watering, G., Gijbels, D., Dochy, F., & Rijit, J. (2008). *Students' assessment preferences, perceptions of assessment and their relationships to study results*. Springer, *56*(6).

Van der Vleuten, C. P. M., Sluijsmans, D., & Joosten-ten Brinke, D. (2017). *Competence assessment as learner support in education*. In Mulder, I. M. (Ed.). *Competence-based vocational and professional education. Bridging the worlds of work and education* (pp. 607–630). Cham, Switzerland: Springer International Publishing Switzerland.

Watkins, D., Dahlin, B., & Ekholm, M. (2005). *Awareness of the backwash effect of assessment: A phenomenographic study of the views of Hong Kong and Swedish lecturers*. Instructional Science, *33*, 283–309.

Wood, R. (1993). *Assessment and testing*. Cambridge: Cambridge University Press.