Teaching English Language Journal
ISSN: 2538-5488 – E-ISSN: 2538-547X – http://tel.journal.org
© 2025 – Published by Teaching English Language and Literature Society of Iran



Please cite this paper as follows:
Khojaste Mehr, M., Mohammadi, M., & Ghanbar, H. (2026). Unveiling Student Teachers' Conceptions of Assessment: The Roles of Gender and Level of Education. *Teaching English Language, 20*(2), 469-508. https://doi.org/10.22132/tel.2024.465139.1647

Research Paper

Unveiling Student Teachers' Conceptions of Assessment: The Roles of Gender and Level of Education

Mohammad Khojaste Mehr
Ph.D. Candidate, Department of English Language Teaching, West Tehran Branch, Islamic Azad university, Tehran
Mojtaba Mohammadi [1]
Assistant Professor, Department of English Language Teaching, West Tehran Branch, Islamic Azad university, Tehran
Hessameddin Ghanbar
Assistant Professor, English Department, Fereshtegaan International Branch, Islamic Azad university, Tehran

Abstract

With the paradigm shift and the reconceptualization of language assessment, language teachers are expected to be --well aware of the concept and its significance in the process of language learning and teaching. This nationwide study aimed to investigate the conceptions of assessment among undergraduate student teachers in TEFL and examine the roles of gender and year of education in shaping these conceptions. The study involved 464 undergraduate student teachers from 12 Farhangian Teacher Training University centers, selected through a convenient sampling method. Participants completed the Students' Conceptions of Assessment Inventory (SCoA-VI), and the results of exploratory factor analysis yielded four conceptions: improvement, affect, student accountability, and teacher

[1] Corresponding author: mojtabamohammadi@gmail.com; mohammadi.mojtaba@wtbiau.ac.ir

accountability. The conception of assessment for improvement had the highest factor loading while teacher accountability had the lowest following affective, and student accountability, respectively. There were no negative conceptions of assessment among the participants and they held more favorable and adaptable conceptions of assessment. Moreover, the results of multivariate analysis of variance did not reveal any significant differences in the conceptions of assessment in terms of gender and year of education. These results have important implications for teacher training courses and higher education centers.

Keywords: Conceptions of assessment; Gender; Student teachers; Teacher training university

Received: June 28, 2024
Accepted: November 04, 2024



1. Introduction

As a result of the recent shift to a sociocultural approach in language education, assessment has gained agency by becoming "a watchword of the twenty-first century" (DeLuca et al., 2024, p. 1) in schools and societies and turning out to be an essential component of any curriculum in educational systems (Alshammari, 2020; Berry, 2011). The main purpose of assessment in an educational system is to identify student's potential for growth and to create a positive difference at an individual and social level through their learning. According to Khan (2012), if educational principals and teachers are seeking an exciting and beneficial educational system, they should engage learners in pedagogical activities, and only in this way can they reveal their ultimate potencies. As one of these activities, the assessment process is of paramount importance for teachers in identifying the learners' strengths and weaknesses, and it greatly impacts students' learning and motivation (McMillan, 2015). Recent research by Coombs and DeLuca (2024) has highlighted the multifaceted nature of assessment discourses, emphasizing

the importance of understanding assessment competence, literacy, capability, and identity among educators. These constructs are crucial for both pre-service and in-service teachers as they navigate the complexities of educational assessment. By mapping these discourses, scholars have provided a comprehensive framework that can inform teacher education programs and enhance the assessment literacy of future educators.

In addition, many studies conducted by different scholars showed that language assessment knowledge was perceived as essential by the teachers (Tavassoli & Farhady, 2018) and had a significant impact on their student's educational success, motivational beliefs, activities in the teaching classroom, and self-regulation skills (Alkharusi, 2013; Dorman et al., 2006; Heritage, 2007; Khojaste Mehr et al., 2024; Yıldızlı & Saban, 2016). Therefore, "educators' perspectives, attitudes on teaching, tendencies, views about the teaching-learning process, and skills and awareness about assessment play a significant role in classroom assessment practices" (Abell & Siegel, 2011, p. 205). Xu et al., (2023) found that EFL teachers' conceptions of assessment are significantly affected by "personal factors (i.e., teachers' agency in assessment), experiential factors [i.e., (anti-)apprenticeship of observation about assessment], and contextual factors (i.e., exam-oriented culture)" (p. 237). Therefore, Given the agentive role of teachers in assessment, language educational centers must prioritize understanding teachers' conceptions of assessment. Therefore, it is deemed necessary to integrate significant reforms in language assessment into teacher education programs since a quality teacher education program is the prerequisite to "tailor the complex and diverse needs of 21st-century learners" (Arifani et al., 2024, p. 1). By doing so, investigating teachers' conceptions of assessment and their practices can play pivotal roles in the learning outcomes of any educational system. This urgency is not overlooked in the Iranian educational context in general, yet

no studies have addressed teacher education universities in particular. Furthermore, no studies have been devoted to investigating these universities' undergraduate student-teachers of the Teaching English as a Foreign Language (TEFL) program.

2. Literature Review

The literature defines conception as the overall, typically implied, understanding an individual has about a phenomenon (Brown, 2008). Conceptions denote the concepts, standards, and approaches individuals have concerning what something is and what its purpose is. Conceptions are shaped progressively through involvement with a phenomenon and develop the system through which an individual's responses or reactions to the phenomenon are made (Fives & Buehl, 2012). Assessment conceptions are teachers' knowledge of the description and goal of examining, testing, and evaluating students' learning. As argued by Segers and Tillema (2011), "studying conceptions of assessment is of utmost relevance at a time when innovation of assessment practces is on the educational agenda" (p. 53). Previous research showed that assessment conceptions are associated with academic performance as a result of improving self-regulation (Brown, 2011). Supporting assessment as a medium for development seems to be a practical self-regulating idea resulting in enhanced academic achievement as self-regulation needs thinking on the accomplished performance to recognize educational primacies and achievements. Gulikers et al. (2008) also pointed out that learners worked harder and established more specialized competencies when they considered assessment an authentic practice. On the other hand, followers of assessment as irrelevant and ignorable mostly reject the validity of the evaluative process, results, or feedback. Brown (2008) believed that all assessment procedures follow three main purposes and one "anti-purpose". These are: (1) it brings about improvement in teaching and

learning, (2) it leads to schools and teachers' accountability, (3) Students are responsible for their learning through assessment, and (4) it is irrelevant to the life and work of teachers and students (irrelevant). The literature has shown that these purposes and anti-purposes differ as a result of context (Farangi & Rashidi, 2022; He, et al., 2011).

2.1 Conceptions of Assessment: Theoretical Background

The idea that conceptions and beliefs have a causal relationship with actions has been riveting for the philosophers and theoreticians who have proposed theories and models to conceptualize the relationship. These theories and models have identified the significance of beliefs and conceptions (in this study language teachers' conceptions of the assessment) as crucial in delving into their thoughts and actions. In the Belief-Desire-Intention (BDI) Model proposed by Bratman (1987), he theorized that our beliefs and desires together with our intentions can help characterize our actions. For him, the first two are fundamental for the existence of the action yet the third is a requirement for the consistency of the action. In Ramsey's Theory of Belief, initially proposed in 1929, he asserted that any person's "actions are indispensable in explaining the notion of belief" (Grubber, 2022, p. 3). In the Theory of Reasoned Action (TRA) by Ajzen (1991), and its successor, the Theory of Planned Behavior (TPB), it is believed that attitudes and beliefs one has toward performing a behavior and action can predict its actual performance. According to the theory, the prediction will be stronger if one's attitudes and conceptions are specific to the expected action and behavior. In the context of this study, understanding the teachers' conceptions of assessment can mainly contribute to the teachers' actual performance in the assessment and its related classroom practices. Furthermore, conceptions of assessment have an impact on the learners' actions in getting ready for assessment, acting through the assessment, and post-evaluation practices of

assessment consequences (Peterson & Irving, 2008). Also, the complex interplay between teachers' and students' conceptions of classroom assessment and their approaches to grading is crucial. Baidoo-Anu et al. (2023) emphasized that understanding these conceptions is necessary for developing effective assessment practices that align with educational goals. This aligns with the findings of the current study, which also underscores the importance of favorable and adaptable conceptions of assessment among student teachers. One more point is that communities expect success from every attempt made by their individuals, and these successes are to be found in the assessment results (Brown, 2022).

Under the sociocultural paradigm, assessment is narrowly connected both to instruction and learning. With such an influence, assessment warrants more devotion from instructors and learners themselves (Pastore & Pentassuglia, 2016). To create responsible, self-reliant learners, educators recommend that students actively participate in the assessment process to understand expectations, including the standards and criteria that must be met and the best ways to achieve desired outcomes. (Rust et al., 2005). Assessment for Learning is described as a rich process with formal and informal feedback that offers prospects to examine and use skills, knowledge, and understandings, involved in authentic assessment tasks, raise learners' independence and autonomy, and provide harmony between summative and formative assessment (Sambell et al. 2012).

Assessment has an authoritative impression on pupils' lives since it adjusts their education procedures and enlightens upcoming choices on several aspects such as occupation (Chen & Brown, 2018). Given the benefits of assessment, it is likely that learners will adjust their belief structures about assessment in response to its associated consequences (Brown et al., 2009).

Brown et al. (2009) contended that assessment can implement adaptive and maladaptive forces on pupils' education. Assessment plays an adaptive part if learners exploit the assessment evidence to categorize their educational requirements and take suitable activities to remediate the flaws. It turns into maladaptive when learners consider assessments as a medium that can be overlooked (Brown et al., 2009). As a result, several conceptions of assessment practices have been proposed in the literature.

2.2 Empirical Studies

Reviewing the literature reveals that teachers have diverse conceptions of assessment and how it can be practiced in the classroom. Some scholars have targeted the students' conceptions and practices while others have focused on the teachers. Matos et al. (2015) found that students held different beliefs, the most common being a negative emotional reaction to accountability. Matos and Brown (2015) further compared university students in Brazil and New Zealand. They found that the former had a more unfavorable view of assessment, which they attributed to the widely summative approach to assessment in universities. Exploring students' perceptions, Brown and Wang (2016) and Landim et al. (2015) used a "draw-a-picture" method to reveal that negative emotions and the perception of assessment as inaccurate or imprecise were prevalent among Brazilian university students, accounting for 60% of the assessment-related drawings. Therefore, although the number of studies on conceptions of assessment is extensive, there are still gaps in the conceptions of assessment studies targeting various groups of teachers and learners especially when considering it along with the effects of interventional practices.

Teachers have also been investigated for their conceptions and practices of assessment. In a study by Yetkin and Özer (2020), conceptions were

recognized as “improvement, school accountability, student accountability, and irrelevance” (p. 384). Ninety-seven student teachers’ results showed that improvement was the most favorable and school accountability was the least favorable conception of assessment. In addition, the conceptions of improvement, school, and student accountability had positive intercorrelations, and gender and grade level/teaching significantly differed in participants’ conceptions. One more study on student teachers conducted by Lutovac and Flores (2021) contended that conceptions of assessment significantly affect future teachers’ pedagogical practices. They disclosed four main conceptions: "(a) assessment is about feedback and reflection, (b) assessment needs to be personalized and demonstrate students’ learning, (c) assessment should take into account students’ invested effort, and (d) assessment fails to measure students’ success or failure" (p. 11). In the context of general education at the tertiary level, Brown (2022) posited that universities assess and evaluate students with respect to their proficiency in essential disciplinary knowledge and skills, which can have an impact on their attitudes, beliefs, and emotions. However, the negative impacts of assessments can also be mitigated “if students regulate their responses to assessment and feedback” (p. 116). According to Pereira et al. (2022), students have their own unique perspectives on the utility of assessments and analyze these perspectives as they participate in the assessment process. The study revealed that the scale used was both valid and reliable, as determined by the factor analysis of its internal structure. Notably, unlike those who preferred traditional methods, those who followed alternative assessment methods tended to score higher in terms of effectiveness, fairness, and engagement with assessment. Barnes et al. (2023) conducted a comprehensive review of teachers’ beliefs about assessment, revealing that these beliefs significantly influence their assessment practices and decision-

making processes. They argue that understanding these beliefs is essential for designing professional development programs that effectively address teachers' needs and improve assessment literacy

The conceptions of assessment are integral in universities since they can build perceptions that may differ from those in compulsory schooling. Higher education may result in having students with various conceptions of assessment due to the widespread use of summative evaluation practices and the fact that students are legally adults and can make their own choices about assessment to achieve their life goals, as noted by Panadero et al. (2020) and Harris et al. (2018). Psychology of assessment (i.e. beliefs and attitudes towards assessment) among student teachers may play an important role in their future careers. Teacher education careers may shape student teachers' conceptions or perceptions in a way that is in line with the top-down procedures and mostly take a summative look (Panadero et al., 2020) while the student teachers should be given enough space to exercise their autonomy in assessment (Harris et al., 2018). Panadero et al. (2020), further, believed that assessment has very important political consequences which can lead to top-down and bottom-up changes in schools and national policies as well as challenge the stakeholders and principals. This high rank in education cannot be ignored in the teacher education programs associated with preservice teachers.

Although some literature is available on learners' conceptions of assessment; student teachers' conception of assessment has garnered scant attention in the SLA literature (Lutovac & Flores, 2021) despite the fact that student teachers are being prepared to start their jobs as teachers and be involved directly with the processes and procedures of assessments. In Iran, Farhangian University is responsible for preparing teachers to start their teaching careers as employees of the Ministry of Education. They are specific

universities that accept students with very good grades in the National Entrance Exam. Therefore, examining their students' conceptions of assessment can add to the previous literature.

On this account, the present study investigated the following research questions:

1. What are the Iranian TEFL student teachers' conceptions of assessment and the underlying components?
2. Is there any significant difference between male and female student teachers on the conceptions of assessment and the underlying components?
3. Is there a statistically significant difference between years of education on the conceptions of assessment and its underlying components for Iranian undergraduate TEFL student teachers?

3. Methodology

3.1 Participants and Setting

Four hundred sixty-four university students participated in the present study. They were both male (n= 331) and female (n= 269) university students of B.A. level majoring in teaching English as a Foreign Language (TEFL) in different semesters from freshman and sophomore to junior and senior; yet, more than half of them had passed the Fundamentals of Assessment as the core course in the TEFL program. The participants were selected from different centers of Farhangian University, a teacher training university, located all over Iran. Their age was between 18 and 29 years old (m= 29) and their native language was Persian. All the participants in the present study consented to participate in this study. Many more students were sent the questionnaire yet they refused to reply. Those students who want to be school

teachers should compete in a national entrance exam at the end of their high school years. They will be trained for four years in this university and then, they will be able to start their work as teachers in the Ministry of Education.

3.2 Instrumentation

The Students' Conceptions of Assessment Inventory (SCoA-VI) developed by Brown, Peterson, and Irving (2009) was used to collect the learners' conceptions of assessment. The original inventory administered in this study consisted of 45 items and four inter-correlated constructs (Improvement: Assessment Increases Learning and Instruction; External: Assessment deals with External Factors; Affect: Assessment has Emotional Profit, and Irrelevance: Assessment is Irrelevant and does not contribute meaningfully to the learning process). However, the components were renamed as a result of this study since the instrument was re-validated among Iranian TEFL students at Farhangian University. The positively-packed rating scale was used for answers (Lam & Klockars, 1982) with two negative choices (strongly disagree, disagree) and four positive choices (slightly, moderately, mostly, and strongly agree). According to Brown (2004), for instruments that measure belief and attitude, respondents tend to answer positively to the statements. Therefore, it is more likely to elicit variance within responses of a positively-packed rating scale compared to balanced response formats like the Likert scale and can help encompass a broader spectrum of favorable positive attitudes and beliefs. The instrument had a reliability index of .85. It was also re-validated by the student teachers of TEFL within the context of Iran. The results are reported in the results section below. the two groups of students: Freshman and Sophomore (the first group) and Junior and Senior (the second group), on the four conceptions of assessment.

3.3 Data collection and analysis

The Students' Conceptions of Assessment Inventory (SCoA-VI) was distributed among the students after receiving formal permission from the central administration of Farhangian University. Following that, one of the researchers visited the university centers in different provinces of Iran in the Fall semester of 2022. The participants were asked to fill out the questionnaire and he was present in case there was a question. It took around 20 minutes for each participant to complete the questionnaire; the data collection took around three months. Exploratory factor analyses were conducted to validate the instrument and extract the constructs among the Iranian university students of TEFL. Multivariate Analysis of Variance (MANOVA) was also run to investigate any significant difference in the independent variable (i.e., students' conception of assessment) with regard to our target dependent variables of gender and years of study.

4. Results

RQ1. What are the Iranian ELT student teachers' conceptions of assessment and the underlying components?

To investigate the factorial structure of the student teachers' conceptions of the assessment questionnaire, its 45 items, framed as a 6-point Likert scale, were submitted to exploratory factor analysis (EFA) adopting principal axis factoring (PAF) as the factor extraction method and Promax as its rotation method. PAF was selected since it would return a factor structure in which, instead of unique variance and error variance, the common variance was accounted for (Tabachnick & Fidell, 2013). The ultimate goal of this technique is to account for the maximum amount of extracted variance (showing the maximum amount of data in the scale). PAF has been chosen

given that it is also more robust than other extraction methods in different conditions of data structure such as non-normality. Additionally, Promax, as an oblique rotation method, was selected, as we observed moderate correlation among factors in our preliminary EFA, and based on this, we utilized multivariate analysis of variance (MANOVA) in the next stage of our analysis.

Before conducting EFA, the soundness of data to run factor analysis was scrutinized. First, the assumption of normality was verified by checking the skewness values of the items (see Table 1) to be between -2 and +2 (see Table 1) which, in line with Tabachnick and Fidell (2013), was met. Secondly, the Kaiser-Meyer-Olkin measure was carried out to examine the sampling adequacy for EFA (i.e., the sufficiency of the sample size). As revealed in Table 2, KMO was 0.89 which exceeded the suggested value of 0.6 (Field, 2024). Thirdly, Bartlett's test of sphericity was X^2 (406) =1882.87, p = .00 (Table 2), signifying that the magnitudes of correlation coefficients between items were amply large for conducting EFA.

Table 1.

Descriptive Statistics of Students' Conceptions of Assessment (N = 464)

	Minimum	Maximum	Mean	Std. Deviation	Skewness	
	Statistic	Statistic	Statistic	Statistic	Statistic	Std. Error
item1	1	6	4.31	1.31	-1.04	0.11
item2	1	6	4.79	1.24	-1.32	0.11
item3	1	6	2.92	1.62	0.17	0.11
item4	1	6	3.94	1.49	-0.62	0.11
item5	1	6	3.58	1.50	-0.35	0.11
item6	1	5	3.27	1.51	-0.23	0.11
item7	1	6	2.05	1.30	1.02	0.11
item8	1	6	3.87	1.45	-0.65	0.11
item9	1	6	4.18	1.36	-0.76	0.11

item10	1	6	3.64	1.63	-0.31	0.11
item11	1	6	3.11	1.61	0.07	0.11
item12	1	5	3.42	1.34	-0.28	0.11
item13	1	6	3.62	1.44	-0.38	0.11
item14	1	6	3.77	1.42	-0.56	0.11
item15	1	6	4.47	1.23	-0.96	0.11
item16	1	6	4.21	1.25	-0.89	0.11
item17	1	6	3.50	1.46	-0.31	0.11
item18	1	6	3.57	1.52	-0.34	0.11
item19	1	5	3.15	1.46	-0.11	0.11
item20	1	6	4.45	1.28	-0.93	0.11
item21	1	6	4.13	1.36	-0.67	0.11
item22	1	6	4.36	1.26	-0.92	0.11
item23	1	6	3.04	1.54	0.05	0.11
item24	1	6	4.22	1.28	-0.96	0.11
item25	1	5	3.14	1.39	-0.10	0.11
item26	1	6	4.21	1.32	-0.74	0.11
item27	1	6	4.14	1.40	-0.66	0.11
item28	1	6	4.23	1.28	-0.77	0.11
item29	1	6	3.78	1.54	-0.51	0.11
item30	1	6	3.06	1.44	0.12	0.11
item31	1	5	3.16	1.67	-0.15	0.11
item32	1	6	3.41	1.42	-0.28	0.11
item33	1	6	3.79	1.38	-0.51	0.11
item34	1	6	3.49	1.46	-0.31	0.11
item35	1	6	3.43	1.41	-0.26	0.11
item36	1	5	3.08	1.59	-0.07	0.11
item37	1	6	4.44	1.10	-0.72	0.11
item38	1	6	3.56	1.44	-0.27	0.11
item39	1	6	3.55	1.51	-0.28	0.11
item40	1	5	3.28	1.58	-0.25	0.11
item41	1	6	3.54	1.40	-0.35	0.11
item42	1	6	4.08	1.24	-0.71	0.11
item43	1	6	2.95	1.48	0.13	0.11
item44	1	5	3.41	1.41	-0.31	0.11
item45	1	6	3.58	1.41	-0.45	0.11

Table 2.

KMO and Bartlett's Test

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.95
Bartlett's Test of Sphericity	Approx. Chi-Square	9787.82
	Df	990
	Sig.	.000

PAF as an extraction method revealed a preliminary ten-factor solution (see Table 3); this solution was obtained using the Kaiser Criterion. We observed however that FIVE factors were represented only by two OR one items, which is quite predictable as we developed this questionnaire by adapting a substantial number of items to fit the new population. According to Meyers et al. (2013) and Keline (2016), for a reliable factor, at least three items are needed. Based on this recommendation, these factors were removed from our analysis given that they were not adequately represented by the items. With regard to the aforementioned precaution, we ended up having a four-factor solution which explained a total of 40.25% of the common variance (see Table 3), with those five factors accounting for 32.29%, 3.06%, 2.57%, and 2.34% of that common variance, correspondingly. It should be mentioned here that items, 10, 29, 17, 34, 1, 3, 4, 36, 19, 40, 44, 31, 6, 25, 12, and 24 were omitted from the factor structure because of either low coefficient (lower than the cut-off value of .47), or loading on unreliable factors.

Table 3.

Total Variance Explained by the Four-factor Solution

Factors	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings
	Total	% of Variance	% Cumulative	Total	% of Variance	Cumulative %	Total
1	**14.96**	**33.25**	**33.25**	**14.53**	**32.29**	**32.29**	**12.38**
2	**1.93**	**4.28**	**37.53**	**1.38**	**3.06**	**35.35**	**10.00**
3	**1.79**	**3.98**	**41.51**	**1.15**	**2.57**	**37.92**	**10.17**
4	**1.58**	**3.51**	**45.02**	**1.05**	**2.34**	**40.25**	**3.49**
5	1.38	3.07	48.09	0.91	2.02	42.27	5.28
6	1.24	2.77	50.86	0.79	1.76	44.02	7.76
7	1.16	2.58	53.44	0.72	1.59	45.62	6.62
8	1.13	2.51	55.95	0.57	1.26	46.87	1.22
9	1.04	2.30	58.25	0.49	1.08	47.96	4.80
10	1.01	2.24	60.48	0.45	1.00	48.96	2.45

Note: Extraction Method: Principal Axis Factoring. Rotation Method: Promax with Kaiser Normalization."

In addition, Table 4, the structure matrix, portrayed the information regarding the extracted factor structure, showcasing the factors, their representing items, and the reliability indices of each factor (all of Cronbach's alphas were larger than the benchmark value of 0.7).

Table 4.

EFA Summary

Construct	Items	Loadings	Cronbach's Alpha
Improvement	Item 21	0.779	0.90
	Item 27	0.733	
	Item 20	0.727	
	Item 16	0.٧22	
	Item 28	0.700	
	Item 26	0.673	
	Item 15	0.650	
	Item 33	0.633	
	Item 42	0.630	
	Item 2	0.597	
Affect	Item 30	0.811	0.89
	Item 35	0.783	
	Item 39	0.756	
	Item 43	0.736	
	Item 18	0.720	
	Item 11	0.696	
	Item 5	0.672	
	Item 38	0.513	
Student Accountability	Item 14	0.739	0.87
	Item 32	0.700	
	Item 13	0.690	
	Item 41	0.617	
	Item 45	0.600	
	Item 8	0.592	
	Item 23	0.586	
Teacher Accountability	Item 37	0.600	0.75
	Item22	0.561	
	Item 9	0.512	

RQ2. Is there any significant difference between male and female student teachers on the conceptions of assessment and the underlying components, as measured by the Students' Conceptions of Assessment (SCoA-VI)?

To answer this question, a two-group multivariate analysis of variance (MANOVA) was employed. More specifically, this MANOVA was used to

examine the effect of gender on four extracted factors/constructs of this questionnaire: (a) improvement, (b) affect, (c) student accountability, and (d) teacher accountability, considered as four main dependent variables (DVs) in this test.

With regard to the measurement of these four factors extracted from the conceptions of the assessment questionnaire, it is worth mentioning that they were all deemed as latent composites; consequently, the means of participants' responses to items of each factor were calculated and used in MANOVA. After estimating those four DVs (factors/constructs), these two groups of students, that is, male and female students (the independent variable [IV] with two levels) were compared with each other to find out whether they were different with regard to various factors of conceptions of assessment. (see Table 5)

Table 5.

Descriptive Statistics of Students' Conceptions of Assessment Constructs across Gender

Construct	Gender	Mean	Std. Deviation	N
Improvement	Female	4.29	0.92	265
	Male	4.19	0.99	199
	Total	4.25	0.95	464
Affect	Female	3.33	1.13	265
	Male	3.38	1.16	199
	Total	3.35	1.14	464
Student Accountability	Female	3.35	0.98	265
	Male	3.38	1.05	199
	Total	3.36	1.01	464

Teacher Accountability	Female	4.37	0.94	265
	Male	4.27	0.90	199
	Total	4.33	0.93	464

Before conducting the two-group MANOVA, we examined the data structure and the key statistical assumptions of the test. We did not find any missing cases, outliers, or influential observations. Relating to MANOVA's statistical assumptions, we did not observe any univariate or multivariate normality violations, given that all skewness and kurtosis values of dependent variables were between -2 and +2 standard errors of their observed values. We also examined the maximum expected values of Mahalanobis distances of DVs. We saw that they were all below their expected values in a chi-square distribution illustrating multivariate normality (Pituch & Stevens, 2016). As with the assumption of homogeneity of covariance matrices, Box's test of homogeneity of covariance matrices did not yield any significant results (this examines the null hypothesis, which states that there is no group difference in the observed covariance matrices of the dependent variables). (see Table 6).

Table 6.

Box's Test of Equality of Covariance Matrices

Box's M	11.209
F	1.110
df1	10
df2	856819.38
Sig.	.350

A two-group MANOVA's findings showed that there is no discernible difference between the two gender groups (i.e. e. pupils

(male and female) based on those four categories related to the conceptions of assessment, that is, improvement, affect, student accountability, and teacher accountability, was not rejected, F Wilk's Lambda (4, 459) = .984, p = .12, partial eta squared of .02 (see Table 7 and Table 8), which can be considered a small effect. Hence, it can be argued that the student's gender did not have any statistically significant holistic effect on their conceptions of assessment.

Table 7.

Multivariate Tests for Investigating the Holistic Effect of Gender on Conceptions of Assessment

	Effect	Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.966	3267.01	4	459	0.00	0.97
	Wilks' Lambda	.034	3267.01	4	459	0.00	0.97
	Hotelling's Trace	28.471	3267.01	4	459	0.00	0.97
	Roy's Largest Root	28.471	3267.01	4	459	0.00	0.97
Gender	Pillai's Trace	.016	1.85	4	459	0.12	0.02
	Wilks' Lambda	**.984**	**1.85**	**4**	**459**	**0.12**	**0.02**
	Hotelling's Trace	.016	1.85	4	459	0.12	0.02
	Roy's Largest Root	.016	1.85	4	459	0.12	0.02

In addition to investigating the holist effect of gender on student teachers' conception of assessment, gender's possible effects as an independent variable on each of the four DVs (the four extracted factors of the questionnaire) were investigated using four univariate F tests (included in the two-group MANOVA). As can be seen in Table 8, gender did not have any significant effect on the four conceptions of assessment. This showed there was no statistically significant difference between the two gender groups on the four conceptions of assessment.

Table 8.

Test of Between-subjects Effects of Gender on Conceptions of Assessment

IV	DVs	Type III Sum of Squares	Df	Mean Square	F	Sig.	Partial Eta Squared
Gender	Improvement	1.10	1	1.10	1.22	0.27	.003
	Affect	0.38	1	0.38	0.29	0.59	.001
	Student Accountability	0.12	1	0.12	0.12	0.73	.000
	Teacher Accountability	1.07	1	1.07	1.25	0.26	.003
Error	Improvement	415.68	462	0.90			
	Affect	601.85	462	1.30			
	Student Accountability	469.30	462	1.02			
	Teacher Accountability	395.46	462	0.86			
Total	Improvement	8796.08	464				
	Affect	5811.81	464				

Student Accountability	5707.53	464
Teacher Accountability	9083.44	464

RQ3. Is there a statistically significant difference between years of education on the conceptions of assessment and the underlying components for Iranian undergraduate TEFL student teachers?

Another two-group multivariate analysis of variance (MANOVA) was adopted to address this question. This was used to examine the effect of the year of education (Freshman, Sophomore, Junior, Senior) on four factors extracted from the questionnaire: (a) improvement, (b) affect, (c) student accountability, and (d) teacher accountability, considered as the four main dependent variables (DVs) in this test.

With regard to the measurement of these four factors of the conceptions of the assessment questionnaire, it is worth mentioning that they were all deemed as latent composites; hence, the means of participants' responses to items of each factor were calculated and used in MANOVA. After estimating those four DVs (factors/constructs), these two groups of students, that is, freshmen and sophomores as the first group and the juniors and seniors as the second group (the independent variables with two levels), were compared to find out whether they were different in terms of various conceptions of assessment (see Table 9 for more information on participants' responses in different related constructs).

Table 9.

Descriptive Statistics of Conceptions of Assessment Constructs across the Year of Education

Construct	Year of Education	Mean	Std. Deviation	N
Improvement	1&2	4.24	0.97	238
	3&4	4.26	0.93	226
	Total	4.25	0.95	464
Affect	1&2	3.27	1.19	238
	3&4	3.44	1.08	226
	Total	3.35	1.14	464
Student Accountability	1&2	3.29	1.00	238
	3&4	3.44	1.01	226
	Total	3.36	1.01	464
Teacher Accountability	1&2	4.37	0.90	238
	3&4	4.28	0.95	226
	Total	4.33	0.93	464

Note: 1&2 = Freshman and Sophomore (the first group), 3&4 = Junior and Senior (the second group)

Before conducting the two-group MANOVA, we scrutinized the data structure and key statistical assumptions of this test. We did not see any missing cases, outliers, and also influential observations. Apropos of MANOVA's statistical assumptions, we did not find any violations of univariate or multivariate normality as all skewness and kurtosis values of dependent variables were between -2 and +2 standard errors of their observed values. We also inspected the maximum expected values of Mahalanobis distances of DVs and observed that they were all less than their expected values in a chi-square distribution which illustrated the multivariate normality (Pituch & Stevens, 2016). Pertaining to the assumption of

homogeneity of covariance matrices, Box's test of homogeneity of covariance matrices did not return any significant results (this tests the null hypothesis that the observed covariance matrices of the dependent variables are equal across groups. (see Table 10).

Table 10.

Box's Test of Equality of Covariance Matrices

Box's M	11.701
F	1.159
df1	10
df2	1014255.86
Sig.	.313

A two-group MANOVA showed that there is no discernible difference between the two groups — Freshman and Sophomore (the first group) as well as Junior and Senior (the second group) — on those four conceptions of assessment, that is, improvement, affect, student accountability, and teacher accountability, was not rejected, F Wilk's Lambda (4, 459) = .979, p = .05, partial eta squared of .02 (see Table 11 and Table 12), which can be considered a small effect. Thus, it can be argued that the length of study of students did not have any statistically significant holistic effect on their conceptions of assessment.

Table 11.

Multivariate Tests for Investigating the Holistic Effect of Year of Education on Conceptions of Assessment

Effect		Value	F	Hypothesis df	Error df	Sig.	Partial Eta Squared
Intercept	Pillai's Trace	.967	3331.65	4	459	0.00	.97

	Wilks' Lambda	.033	3331.65	4	459	0.00	.97
	Hotelling's Trace	29.034	3331.65	4	459	0.00	.97
	Roy's Largest Root	29.034	3331.65	4	459	0.00	.97
Year of education	Pillai's Trace	.021	2.421	4	459	.05	.02
	Wilks' Lambda	**.979**	**2.421**	**4**	**459**	**.05**	**.02**
	Hotelling's Trace	.021	2.421	4	459	.05	.02
	Roy's Largest Root	.021	2.421	4	459	.05	.02

In addition to investigating the holist effect of the year of education on students' conceptions of assessment, the potential effects of the year of education as an independent variable on each of the four DVs (the four extracted factors of the questionnaire) were investigated separately using four univariate F tests (included in the two-group MANOVA). As it was shown in Table 12, the year of education did not have any significant effect on the four conceptions of assessment. This indicated that there was not any statistically significant difference between the two groups of students: Freshman and Sophomore (the first group) and Junior and Senior (the second group), on the four conceptions of assessment.

Table 12.

Test of Between-subjects Effects of Year of Education on Conceptions of Assessment

Source	Type III Sum of Squares	df	Mean Square	F	Sig.	Partial Eta Squared

Year of education	Improvement	.03	1	.031	.034	.853	.000
	Affect	3.27	1	3.273	2.52	.113	.005
	Student Accountability	2.73	1	2.728	2.70	.101	.006
	Teacher Accountability	1.02	1	1.020	1.19	.276	.003
Error	Improvement	416.75	462	.902			
	Affect	598.95	462	1.296			
	Student Accountability	466.70	462	1.010			
	Teacher Accountability	395.52	462	.856			
Total	Improvement	8796.08	464				
	Affect	5811.81	464				
	Student Accountability	5707.53	464				
	Teacher Accountability	9083.44	464				

5. Discussion

In the final model of the present study, as shown in Table 4 above, four factors represented conceptions of assessment among student teachers in a teacher training university. The participants of the present study had four conceptions of assessment: Improvement (assessment improves learning and teaching in students and teachers respectively), Affect (assessment can positively affect the social and emotional aspects of learning), Student accountability (assessment can ensure the quality of students learning), and Teacher accountability (assessment ensures the school quality). Improvement had the highest factor loading while teacher accountability had the lowest. Brown et al.'s (2009, 2016) studies indicated that students who are more successful academically accept assessment as a tool for improvement. This claim is confirmed in the present study since the students who enter Farhangian university centers all over Iran have higher average scores and

they rank higher in the nationwide entrance exam which is held every year for the candidates. This way, and based on the previous studies, they are expected to have more adaptive conceptions towards assessment. The results of this study showed that student teachers consider assessment as a means for both their own and their teachers' improvement (Brown et al., 2009, 2016; Flores et al., 2020). The results also suggest that student teachers do not have negative conceptions of assessment which can encourage teachers to stop "resisting or subverting assessment policies and intended practices" (Deneen & Brown, 2016, p. 2).

The result of this study is in line with that of Remesal and Brown (2015) in which the participants agreed that the main purpose of assessment is improvement as a result of comparing teachers' conceptions of assessment in compulsory and non-compulsory contexts. They also believed that context plays an important role in shaping students' conceptions and their results as those teachers who worked in a compulsory formal setting tended to show a more favorable attitude towards assessment. The participants of this study were student teachers who were studying in a compulsory formal setting with direct consequences for their occupations. As prospective teachers, they take the assessment as one of the main tools in their hands and have a favorable attitude toward it which confirms previous studies (Blazar and Kraft,2017; Brown et al., 2009, 2016; Flores et al., 2020).

These results were also consistent with some other studies and inconsistent with others. Using a new self-report scale, Brown and Wang (2016) compared Hong Kong and Chinese university students' conceptions of assessment and found that assessment as indicative of school quality was the central factor in their study with 7 other factors which hinged on this (Societal Uses, Class Benefits, Accuracy, Negative Aspects, Teacher Use,

Family Effects, and Competition). These factors were different from the factors of the present study which can be explained with regard to the fact that there are cultural differences between the two contexts and also the scales used in these studies were different. Portuguese university students' conceptions of assessment were also investigated during the time of reform in educational procedures by Flores et al. (2020). The same inventory (SCoA-VI) was also used and the researchers reported eight factors including assessment "predicts student future; evaluates school quality; supports student learning improvement; supports the improvement of teaching; is bad; is ignored; is personally enjoyable; and supports positive class environment" (p. 389). Similar to the current study, improvement was the main assessment conception among university students.

Similar to this study, Michaelides and Solomonidou (2019) also found a four-factor model of conceptions of assessment among 7th to 9th-grade students in Cyprus. They used the Students' Conceptions of Assessment-VI inventory and found the following factors: "improvement of teaching and learning, external motives, affective/social activity, and negative, unfair, or valueless practice" (p. 1). Improvement and affective factors were also reported in the current study while external motives and assessment as a negative practice were not. These differences in the results may refer back to the differences in the population of the two studies as those in the current study were student teachers who were intended to be teachers in the future while those in the Michaelides and Solomonidou (2019) study were school students who generally don't consider assessment as their favorite procedure. Hirschfeld and Brown (2009) argued that students who take a more positive view towards assessment are more likely to receive higher grades and this fact is in line with the conceptions of the participants in this study as those

who enter Farhangian Teacher Training University centers in Iran are among the best students in high schools due to the fact that they will be employed right after their graduation. Therefore, the competition for entering Farhangian Universities in Iran is very high among Iranian students and those who pass the highest criteria can enter it. As a result, the researchers can explain their findings in this regard. We expected that the maladaptive conception of assessment would be very low among these students and this was verified in the present study.

Lutovac and Flores (2021) used student teachers' narratives as a medium through which they could delve into their conceptions of assessment. In fact, through the use of a qualitative method, they found out that for student teachers, assessment for improvement and assessment for student accountability were the main conceptions. Although we adopted a quantitative method in this study, we were able to extract similar conceptions from our participants. These results can show that the conceptions held by students can be similar worldwide with slight variations due to the context. Students of secondary schools in Solomonidoua and Michaelides's (2016) study also considered assessment as a tool for improvement. This point is noteworthy as those students who take a positive attitude toward assessment in their school years may bring that attitude to university years and then they may be able to practice that attitude in their teaching careers if they choose to do so. Brown et al. (2014) study which was conducted in the Iranian context among Iranian university students showed that although students conceptualized assessment as a tool for improvement, they also held that assessment is bad for learning. Farangi and Rashidi (2022) conducted the same search this time among Iranian EFL teachers and found a negative factor for assessment. The difference between the previous two studies and the current one is that in the current study, the researchers did not find a

negative component related to the conceptions of assessment which can be considered a new finding in this particular context. Also, it can show that those students who choose teaching as their career and enter Farhangian University tend to take a more positive approach toward assessment. According to Zolfaghari et al. (2022), both teachers and students have shown positive attitudes toward learner-centered pedagogy and assessment practices in teacher training programs at Iranian universities

The results of the second research question were indicative of the fact that male and female student teachers did not hold different conceptions of assessment These results can be explained considering the fact that the levels of academic achievement of these students do not differ a lot and that all of them were among those who ranked very high in their exams, their attitude towards assessment is favorable and the majority of the take assessment as a tool for improvement. These results also differed from some of the previous literature. While the current study did not find any significant differences between male and female student teachers, Michaelides and Solomonidou (2019) showed that female students had a more positive (adaptive) attitude toward assessment than male students. In that study, female students considered “assessment as a means for improving their learning and as more useful and relevant” (p. 6). Therefore, it can be argued that the results of this study were similar to those of female participants in Michaelides and Solomonidou's (2019) study but not to the male participants. Yetkin and Özer's (2020) study explored conceptions of assessment among preservice English teachers in Turkey and the results showed that gender was not significantly related to the conceptions of teachers which is similar to the results of the present study.

The results of the third research question can be explained in the light of the fact that as Farhangian University is a centralized top-down university and all materials are dictated by the above authorities with less freedom for teacher educators, similar courses or materials are taught to these students during their study time and this may not have the potential to instill different conceptions among student teachers. Although the results in the current study showed that there were no significant differences between the year of education and conceptions of assessment, Michaelides and Solomonidou (2019) showed that there was a small significant negative correlation between the school year and conceptions of assessment as the students in the beginning years of their school hold a more positive attitude towards assessment. Panadero et al. (2023) similarly showed that there were no significant differences between students' year level and their self-assessment practices and beliefs which is in line with the results of this study while their previous study Panadero et al. (2020) showed significant differences between years of education and assessment practices which is not in line with the results of this study. These differences can be explained related to the fact that the participants of the two studies differed. While the participants of Panadero et al.' (2020) study were secondary school students and they showed no significant differences, the participants of Panadero et al.'s (2023) study were university students. Therefore, these results were similar to the results of the current study. Moreover, similar to the results of the present study, the Flores et al. (2020) study did not show a significant difference between the length of study and conceptions of assessment among Portuguese university students. Yetkin and Özer's (2020) study produced similar results as they did not find a relationship between years of education and conceptions of assessment among preservice English teachers in Turkey.

6. Conclusion

Assessment is an indispensable component of teacher education programs whose agentic role has been self-evident. The current study examined conceptions of assessment among Iranian student teachers. The large sample was drawn from a specific context of Farhangian teacher training university centers in different provinces in Iran. The results indicated that student teachers have four conceptions of improvement, affect, student, and teacher accountability. The internal conception that assessment improves learning received the most endorsement and the conception of teacher accountability, as an external factor, received the weakest endorsement. The participants revealed a positive attitude toward assessment which can be a characteristic feature of this specific cohort of student teachers. Gender and level of education were not significantly related to the components of conceptions of assessment. The findings of this study were restricted to the quantitative method which could be enriched with qualitative data. Also, the analysis was based on data collected from limited centers of Farhangian University and could be more representative if the sample were selected from centers in all parts of Iran. One implication of this study is the fact that student teachers who are being prepared to teach in formal school contexts may have different conceptions of assessment compared to other stakeholders like policymakers, administrators, and even parents. Also, policymakers are expected to make amendments to the current policies in the Iranian Teacher education programs so that the roles of student and teacher accountability in language assessment are perceived more strongly and positively. Given the students' positive perceptions of assessment *as a means of learning* and *a way to bring them personal enjoyment*, it is more likely that teacher training programs in Iran are moving in a more appropriate direction compared to a few decades ago.

References

Abell, S. K., & Siegel, M. A. (2011). Assessment literacy: What science teachers need to know and be able to do? In D. Corrigan, J. Dillon, & R. Gunstone (Eds.). *The professional knowledge base of science teaching* (pp. 205–221). The Netherlands: Springer.

Ajzen, I., Fishbein, M. (1980). *Understanding attitudes and predicting social behaviour*. Prentice Hall.

Alkharusi, H. (2013). Canonical correlational models of students' perceptions of assessment tasks, motivational orientations, and learning strategies. *International Journal of Instruction*, *6*(1), 21-38.

Alshammari, W. (2020). *The predictive validity of the admission standards in the college of education at King Faisal University in Saudi Arabia* [Doctoral dissertation, University of Southampton].

Arifani, Y., Megawati, F. & Mulyadi, D. (2024). Editorial: Strengthening the quality of teacher education programs. *Frontiers in Education, 8*. https://doi.org/10.3389/feduc.2023.1338688

Baidoo-Anu, D., Rasooli, A., DeLuca, C., & Cheng, L. (2023). Conceptions of classroom assessment and approaches to grading: Teachers' and students' perspectives. *Education Inquiry.* https://doi.org/10.1080/20004508.2023.2244136

Berry, R. (2011). Assessment Reforms Around the World. In R. Berry & B. Adamson (Eds.), *Assessment reform in education. Education in the Asia-Pacific region: Issues, concerns, and prospects (Vol. 14)* (pp 89–102). Springer, Dordrecht. https://doi.org/10.1007/978-94-007-0729-0_7

Blazar, D., & Kraft, M. A. (2017). Teacher and teaching effects on students' attitudes and behaviors. *Educational Evaluation and Policy Analysis, 39*(1), 146-170. https://doi.org/10.3102/0162373716670260

Bratman, M. E. (1987). *Intentions, plans, and practical reason*. Harvard University Press.

Brown, G. T. L. (2022). Student conceptions of assessment: Regulatory responses to our practices. *ECNU Review of Education*, *5*(1), 116-139. https://doi.org/10.1177/20965311211007869

Brown, G. T. L., Peterson, E. R., & Irving, S. E. (2009). Beliefs that make a difference: Adaptive and maladaptive self-regulation in students' conceptions of assessment. In D. M. McInerney, G. T. L. Brown, & G. A. D. Liem (Eds.), *Student perspectives on assessment: What students can tell us about assessment for learning* (pp. 159–186). Information Age Publishing.

Brown, G. T. L., Pishghadam, R., & Shayesteh Sadafian, S. (2014). Iranian university students' conceptions of assessment: Using assessment to self-improve. *Assessment Matters, 6*, 5-33.

Brown, G. T. L., & Wang, Z. (2016). Understanding university student conceptions of assessment: Cultural similarities and jurisdictional differences between Hong Kong and China. *Social Psychology of Education, 19*, 151–173.

Chen, J., & Brown, G. T. L. (2018). Chinese secondary school students' conceptions of assessment and achievement emotions: Endorsed purposes lead to positive and negative feelings. *Asia Pacific Journal of Education, 38*(1), 91-109. https://doi.org/10.1080/02188791.2018.1423951

Coombs, A., & DeLuca, C. (2024). Mapping the constellation of assessment discourses: A scoping review study on assessment competence, literacy, capability, and identity. *Educational Assessment, Evaluation and Accountability*. https://doi.org/10.1007/s11092-022-09389-9

DeLuca, C., Willis, J., Cowie, B., Harrison, C., & Coombs, A. (2024). *Learning to assess: Cultivating assessment capacity in teacher education*. Springer Nature.

Deneen, C. C., & Brown, G. T. (2016). The impact of conceptions of assessment on assessment literacy in a teacher education program. *Cogent Education*, *3*(1). https://doi.org/10.1080/2331186X.2016.1225380

Dorman, J. P., Fisher, D. L., & Waldrip, B. G. (2006). Classroom environment, students' perceptions of assessment, academic efficacy and attitude to science: a LISREL analysis. In Fisher, D. L. & Khine, M. S. (Eds.), *Contemporary approaches to research on learning environments: Worldviews* (pp. 1–28). Singapore: World Scientific.

Farangi, M. R., & Rashidi, N. (2022). The relationship between Iranian EFL teachers' conceptions of assessment and their self-efficacy. *International Journal of Language Testing*, *12*(2), 59-75. https://doi.org/10.22034/ijlt.2022.157125

Field, A. (2024). *Discovering statistics using IBM SPSS statistics* (6th ed.). Sage publications.

Fives, H., & Buehl, M. M. (2012). Spring cleaning for the "messy" construct of teachers' beliefs: What are they? Which have been examined? What can they tell us? In K. R. Harris, S. Graham, & T. Urdan (Eds.), *APA Educational Psychology Handbook*, Vol 2: Individual differences and

cultural and contextual factors (pp. 471–499). Washington, DC: APA. https://doi.org/10.1037/13274-000

Flores, M. A., Brown, G., Pereira, D., Coutinho, C., Santos, P., & Pinheiro, C. (2020). Portuguese university students' conceptions of assessment: Taking responsibility for achievement. *Higher Education, 79*, 377-394. https://doi.org/10.1007/s10734-019-00415-2

Gruber, M. (2022). Ramsey's theory of belief. *European Journal of Pragmatism and American Philosophy*, *14*(XIV-2).

Gulikers, J. T., Kester, L., Kirschner, P. A., & Bastiaens, T. J. (2008). The effect of practical experience on perceptions of assessment authenticity, study approach, and learning outcomes. *Learning and Instruction*, *18*(2), 172-186.

He, Y., Levin, B. B., & Li, Y. (2011). Comparing the content and sources of the pedagogical beliefs of Chinese and American pre-service teachers. *Journal of Education for Teaching*, *37*, 155–171. https://doi.org/10.1080/02607476.2011.558270

Heritage, M. (2007). Formative assessment: What do teachers need to know and do? *Phi Delta Kappan*, *89*(2), 140-145.

Khan, B. (2012). Relationship between assessment and students' learning. *International Journal of Social Sciences and Education, 2*(1), 576-588.

Khojaste Mehr, M., Mohammadi, M., & Hessameddin, G. (2024). Language Assessment Literacy Instruction: Inquiry-Based vs. Expository Approaches to Enhance Student Teachers' Conceptions of Assessment. *Journal of Language and Translation*, *14* (1), 191-205.

Matos, D. A. S., & Brown, G. T. L. (2015). Comparing university student conceptions of assessment: Brazilian and New Zealand beliefs. *Feedback, identidade, trajetórias escolares: Dinâmicas e consequências*, 177-194.

Landim, M. V., Matos, D. A. S., & Brown, G. T. L. (2015). *A avaliação sob o ponto de vista dos estudantes: O uso de desenhos para a análise de concepções de avaliação* [Assessment from the point of view of students. The use of drawing to analyze conceptions of assessment]. Psicologia, Educação e Cultura, 19(2), 55–74.

Lutovac, S., & Flores, M. A. (2021). Conceptions of assessment in pre-service teachers' narratives of students' failure. *Cambridge Journal of Education*, *52*(1). https://doi.org/10.1080/0305764X.2021.1935736

McMillan, J. (2015). *Principles and practices for in-class evaluation, effective teaching based on effective criteria* (A. Bee, Trans. Ed.). Konya: Education Publishing House.

Michaelides, M. P., & Solomonidou, G. (2019). Factorial structure, gender invariance, and predictive validity of the Students' Conceptions of Assessment-VI Inventory. *European Journal of Psychological Assessment, 35*(2), 248–254. https://doi.org/10.1027/1015 5759/a000383

Panadero, E., Fernández-Ruiz, J., & Sánchez-Iglesias, I. (2020). Secondary education students' self-assessment: The effects of feedback, subject matter, year level, and gender. *Assessment in Education: Principles, Policy & Practice*, *27*(6), 607-634. https://doi.org/10.1080/0969594X.2020.1835823

Panadero, E., Pérez, D.G., Ruiz, J.F, Fraile, J., Sánchez-Iglesias, I., & Brown, G. T. (2023). University students' strategies and criteria during self-assessment: Instructor's feedback, rubrics, and year level effects. *European Journal of Psychology of Education, 38*(3), 1031-1051. https://doi.org/10.1007/s10212-022-00639-4

Pastore, S., & Pentassuglia, M. (2016). Teachers and students' conceptions of assessment within the Italian higher education system'. *Practitioner Research in Higher Education Journal, 10*(1), 109- 120.

Pereira, D., Cadime, I., Brown, G., & Flores, M. A. (2022). How do undergraduates perceive the use of assessment? A study in higher education. *European Journal of Higher Education*, *12*(1), 1-17.

Peterson, E. R., & Irving, S. E. (2008). Secondary school students' conceptions of assessment and feedback. *Learning and Instruction, 18*(3), 238–250. https://doi.org/10.1016/j.learninstruc.2007.05.001

Remesal, A., & Brown, G.T.L. (2015). Conceptions of assessment when the teaching context and learner population matter: compulsory school versus non-compulsory adult education contexts. *European Journal of Psychology Education, 30*, 331–347. https://doi.org/10.1007/s10212-014-0236-3

Rust, C., O'donovan, B., & Price, M. (2005). A social constructivist assessment process model: how the research literature shows us this could be best practice. *Assessment & Evaluation in Higher Education*, *30*(3), 231-240.

Sambell, K., McDowell, L., & Montgomery, C. (2012). *Assessment for learning in higher education*. Routledge.

Scriven, M. (1967). The methodology of evaluation. In R. W. Tyler, R. M. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (Vol. 1, pp. 39–83). Chicago, IL: Rand McNally.

Segers, M., & Tillema, H. (2011). How do Dutch secondary teachers and students conceive the purpose of assessment? *Studies in Educational Evaluation*, *37*(1), 49-54.

Solomonidoua, G., & Michaelides, M. (2016). Students' conceptions of assessment purposes in a low stakes secondary-school context: A mixed methodology approach. *Studies in Educational Evaluation, 52*, 35-41. https://doi.org/10.1016/j.stueduc.2016.12.001

Tavassoli, K., & Farhady, H. (2018). Assessment Knowledge Needs of EFL Teachers. *Teaching English Language, 12*(2), 45-65.

Xu, Y., Chen, J., & He, L. (2023). How personal, experiential, and contextual factors mediate EFL teachers' conceptions of assessment: A narrative study. *Chinese Journal of Applied Linguistics, 46*(2), 237-252. https://doi.org/10.1515/CJAL-2023-0206

Yetkin, R., & Özer, Z. (2020). An investigation into in-service and pre-service English teachers' conception of assessment. *Kuramsal Eğitimbilim Dergisi [Journal of Theoretical Educational Science], 13*(2), 384-396.

Yıldızlı, H., & Saban, A. (2016). The effect of self-regulated learning on sixth-grade Turkish students' mathematics achievements and motivational beliefs. *Cogent Education, 3*(1), 1-17.

Zolfaghari, S., Ashraf, H., Khodabakhshzadeh, H., & Zareian, Gh. (2022). Examining learner-centered pedagogy and assessment practices in

the teacher training program at universities of Iran: Investigating teachers' and students' attitudes. *Teaching English Language, 16*(1), 235-257. https://doi.org/10.22132/TEL.2022.151498

2026 by the authors. Licensee Journal of Teaching English Language (TEL). This is an open access article distributed under the terms and conditions of the Creative Commons Attribution–NonCommercial 4.0 International (CC BY-NC 4.0 license). (http://creativecommons.org/licenses/by-nc/4.0).