

Social Justice in National University Entrance Exams: Insights from EFL Learners' and Teachers' Perceptions in Iran

Sanaz Behboudi¹, Masood Siyyari^{*2}, Gholam Reza Abbasian³

¹Ph.D. Student of Department of Language Teaching, SR. C., Islamic Azad University, Tehran, Iran

²Assistant Professor of Applied Linguistics, Department of Language Teaching, SR. C., Islamic Azad University, Tehran, Iran

³Associate Professor of TEFL, Department of English Language and Literature, Imam Ali University, Tehran, Iran

DOI: [10.71864/LCT-2025-1218019](https://doi.org/10.71864/LCT-2025-1218019)

Received:
14/07/2025

Revised:
18/10/2025

Accepted:
20/10/2025

Published:
18/11/2025

Keywords:

Iranian EFL
Context;
Reliability;
Responsibility;
Social Justice;
Validity

Abstract: Although the influence of social justice on student learning is recognized, there remains a gap in research regarding its significance in language assessment, especially within Iranian English as a Foreign Language (EFL) contexts. As a result, this study explores the perceptions of EFL learners and teachers regarding how well the National University Entrance Exam for Foreign Languages (NUEFL, often referred to as “Konkour”) aligns with principles of social justice. A sequential explanatory mixed-methods design was employed in this study. Data were gathered using a validated social justice questionnaire (SJQ) for quantitative insights, while qualitative perspectives were obtained through semi-structured interviews. Two hundred undergraduate Konkour candidates and 100 English language teachers were selected for the purpose of the present study. In the qualitative phase, 30 teachers and 50 students were selected through purposive sampling from those who responded to the SJQ. The SJQ, comprising 14 items and utilizing a 5-point Likert scale, assessed perceptions regarding social class treatment, individual biases, and the characteristics of test developers. Quantitative data underwent analysis using SPSS, and thematic analysis was employed on qualitative data. The results showed that the SJQ was reliable and had great construct validity. Furthermore, both EFL teachers and learners perceived the Konkour as socially inequitable, citing prejudice towards socioeconomically advantaged groups, the impact of personal bias and subjective assessment, and the uneven application of rights. The qualitative insights further elucidated systemic inequities, including advantages for affluent, urban students, linguistic barriers, and a pervasive distrust in the test's fairness, underscoring that socioeconomic status, geographic location, and resource accessibility substantially influence performance beyond mere ability. The findings underscore ongoing social justice challenges in high-stakes language testing in Iran, stressing the necessity for policy reform, increased teacher awareness, and test development practices that prioritize fairness and equity.

1. Introduction

The notion of fairness and justice has consistently captivated researchers from several fields (Sadeghi et al., 2023). Linguistic theories and cultural conceptualizations are essential in influencing perceptions of

* Corresponding Author's E-mail address: m.siyyari@srbiau.ac.ir



This work is licensed under a [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/).

Cite this article: Behboudi, S., Siyyari, M., & Abbasian, Gh. R. (2025). Social Justice in National University Entrance Exams: Insights from EFL Learners' and Teachers' Perceptions in Iran. *Journal of Language, Culture, and Translation*, 8(1), 254–274. <https://doi.org/10.71864/lct-2025-1218019>

fairness and justice in language teaching and evaluation (Sharifian, 2012). The focus on justice in language assessment has mostly revolved around standardized tests, which are often high-stakes and hence need equitable design and administration (e.g., Parviz, 2023; Song, 2016; Matthew & Coral, 2021). Given that these assessments dictate eligibility for higher education or employment, any manifestation of inequity can yield significant adverse effects for students, including demotivation, test anxiety, and diminished self-efficacy regarding their academic capabilities (Isbell et al., 2023; Karataş & Okan, 2021; Rasooli et al., 2019).

In language testing, fairness pertains to the extent to which test outcomes are accurate, trustworthy, impartial, and uniformly accessible to all examinees, irrespective of their social, cultural, or linguistic backgrounds (Kunnan, 2000, 2004; ETS, 2014). Kunnan's Test Fairness Framework (2000, 2004) stipulates that a fair test guarantees (a) validity and reliability, (b) lack of prejudice, (c) equal access, (d) similar administration, and (e) beneficial societal outcomes. Kane (2010) similarly incorporates fairness into the argument-based validation method, contending that the interpretations and applications of test results must be substantiated by data demonstrating that no grouping is routinely disadvantaged. Davies (2010) asserts that fairness represents an ethical aspect of validity, since assessment judgments possess social and moral ramifications that might impact learners' educational and professional trajectories.

Social justice, as an ancillary concept, pertains to the overarching ethical and societal aspects of assessment, highlighting equitable access to opportunities, acknowledgment of diversity, and the rectification of systemic disparities in evaluation contexts (Novak, 2009; Di Battista et al., 2014; Sadeghi et al., 2023). From this perspective, test fairness functions at the instrumental level (guaranteeing validity and equality in design and scoring), while social justice acts at the societal level (ensuring that testing methods foster inclusiveness and do not perpetuate social stratification).

Institutional and professional norms bolster these ideals. The Educational Testing Service (ETS, 2014) designates fairness as a fundamental component of all testing programs, necessitating equal opportunity for all examinees irrespective of gender, socioeconomic level, or cultural background. The Joint Committee on Testing Practices (2004) defines fairness as including equitable treatment, accessibility, openness, and cultural sensitivity across all phases of test preparation and administration. Recent research (e.g., McNamara & Ryan, 2021; Xi, 2021;

Fan & Kunnan, 2020) underscores that fairness must be defined dynamically, including altering social settings, multilingual realities, and ethical responsibility in evaluation.

This research examines the degree to which social justice principles are integrated into the design of the Konkour in Iran, based on established theoretical and ethical underpinnings. This study specifically investigates whether the Konkour embodies the values of social justice among candidates. Considering that this high-stakes examination profoundly impacts applicants' academic, professional, and personal trajectories (Parviz, 2023) and their educational motivation (Taheri-Larki & Eidy, 2022), it is crucial to ensure it is devoid of bias, prejudice, and social injustice. This study's secondary purpose is to inform test designers and policy-makers of possible deficiencies in the examination, enabling them to create a more equal, legitimate, and socially responsible testing system.

2. Literature Review

2.1. *The Social Dimension of Language Testing*

According to Sadeghi et al. (2023), the concept of fairness and justice has piqued the attention of scientists and researchers working in a variety of sectors. According to Sharifian (2012), linguistic theories and cultural conceptualizations have a substantial impact on how people perceive fairness and justice in the context of language education and assessment. Numerous facets of testing fairness have been discovered and emphasized in the literature (Kunnan, 2000; Shohamy, 2000). Some of these characteristics include item bias, test consequences, score utilization, and fairness linked to standardization. These are only a few of the numerous dimensions of testing fairness.

Language testing is not only a psychometric or technical undertaking; rather, it is a socially embedded practice that is molded by political, cultural, and educational forces. This is something that is underlined in the book *Language Testing: The Social Dimension* (Roever & McNamara, 2006). Exams are not objective instruments; rather, they are a manifestation of the concepts and power dynamics of the groups that are responsible for their creation and implementation. Due to the fact that they control access to educational and professional opportunities, they are responsible for a significant amount of ethical and social responsibilities (Roever & McNamara, 2006).

The pursuit of equity in testing and assessment programs has garnered the interest of researchers including Camilli (2006, 2018), Kunnan (2004, 2005, 2010), Xi (2010), Chan et al. (1998), Takala and Kaftandjieva (2000), Davies (2010), Piller (2021), Ahmadi Safa and Beheshti (2025),

and Beheshti and Ahmadi Safa (2023), among others. In the context of Iran, it is considered that the majority of high-stakes language assessments are unfair since they lack adequate validity (Safari, 2016; Kamyab, 2008). Notwithstanding the fact that NUEEFL has traditionally served as a qualifying tool for admission to Iranian universities, it has not been subjected to substantial revisions for many decades (Naseriazar & Badrian, 2010; Kamyab, 2007). As a consequence of this, there are English as a Foreign Language (EFL) instructors and students who contend that the NUEEFL does not possess construct validity and does not effectively assess the competencies that it purports to examine.

Significant political, social, and ethical repercussions are associated with test and item bias, and these ramifications are felt by L2 test administrators, producers, legislators, and participants. In spite of the fact that there is a noteworthy dearth of rigorous empirical research in language testing that investigates fairness from a social viewpoint (Shohamy, 2000; Kunnan, 2000), this issue has generated a great deal of debate.

2.2. Test Fairness and Social Justice

While fairness and justice are interconnected, they are fundamentally different concepts. McNamara and Ryan (2011) emphasize that fairness relates to the technical characteristics of a test, including its psychometric quality and procedural equality, whereas justice addresses the social implications of test application and the ideological beliefs that support testing frameworks.

2.2.1. Frameworks for Assessing Test Equity

In the realm of educational assessment, three primary frameworks for understanding test fairness have been put forward:

- **Distinction between fairness and validity:** In initial frameworks (Kane, 2010), fairness was viewed as an independent criterion from validity, emphasizing the importance of equitable treatment and the elimination of bias among those taking the test.
- **Fairness that includes various dimensions of validity:** Kunnan (2000) presented a Test Fairness Framework (TFF), which integrates validity, the elimination of bias, accessibility, administration, and social implications. This viewpoint broadens the concept of fairness to encompass not just psychometric precision but also equity and inclusivity throughout the processes of test design, administration, and their subsequent effects.

- **Fairness as essential to validity:** In contrast, Messick (1989), along with AERA et al. (1999) and Davies (2010), regarded fairness as a fundamental aspect of validity. This model illustrates the interconnectedness of fairness and validity, where the evidential and consequential foundations of validity collectively influence the fairness and justice of a test's application and interpretation.

This evolution, from perceiving fairness as distinct from validity to recognizing it as a fundamental aspect of validity, illustrates a conceptual transition from a technocentric perspective to a sociocentric approach to testing, wherein ethical and social considerations are acknowledged as essential to the quality of assessments.

2.2.2. Fairness, Justice, and Social Implications

Messick's (1989) validity matrix offers a structured approach that connects the evidential and consequential aspects of test interpretation and application (see Table 1). Fairness fundamentally pertains to the evidential aspect, encompassing construct validity, reliability, and the absence of bias, while justice broadens the scope to include the consequential aspect, which involves value implications and social outcomes. This distinction resonates with Shohamy's (1997) differentiation between bias inherent in the test (methodological sources) and bias arising from the application of the test (social and political consequences).

Table 1. *Messick's (1989) validity matrix*

	Test Interpretation	Test Use
Evidential Basis	Construct Validity (CV)	CV + Relevance/Utility (R/U)
Consequential Basis	CV + Value Implications (VI)	CV + R/U + VI + Social Consequences

In the realm of language assessment, the emphasis of research has largely been on fairness as opposed to justice. Research has explored the concept of fairness in relation to factors including gender, language background, and content knowledge (Ryan & Bachman, 1992). Nevertheless, there is a scarcity of research focusing on the justice aspect, specifically the implications of value and the social repercussions of test utilization, which align with the third and fourth components of Messick's matrix.

2.2.3. Social Justice in Language Education

Social justice in education entails the establishment of fair conditions that ensure all learners can access opportunities, irrespective of their

social, economic, or cultural backgrounds. Torres (2015) emphasizes the responsibility of educators to facilitate students' engagement with principles of social justice in order to prepare them for future challenges. In a similar vein, Blake (2015) posits that educational institutions have the potential to act as catalysts for social transformation and creative advancement.

Within the context of EFL, Rodas and Osborn (2016) contend that the development of language proficiency occurs concurrently with learners' cognitive integration of social justice concepts. Therefore, language education ought to go beyond mere linguistic skills to encompass critical awareness and a sense of moral responsibility (Hossain, 2018). Considering the varied cultural and social contexts of learners, it is crucial to advocate for social justice in EFL classrooms to enhance inclusion and equity (Hossain, 2018).

2.3. The Theoretical Framework of the Study

This research utilizes Kunnan's (2000, 2004, 2010) Test Fairness Framework (TFF) as its principal theoretical model. The TFF is extensive and multifaceted, including five interconnected components:

1. **Validity:** The extent to which interpretations of test results are supported by facts and theoretical frameworks.
2. **Absence of Bias:** The degree to which items and methods are devoid of discriminatory impacts on subgroups.
3. **Access:** Equitable opportunity for all applicants to exhibit their competencies.
4. **Administration:** Uniform testing circumstances guaranteeing procedural equity.
5. **Social Consequences:** The extensive social and educational ramifications of test use.

This framework aligns the social aspect of testing (Roever & McNamara, 2006) with the larger conversation surrounding fairness in assessments and its connection to social justice (Novak, 2009; Blake, 2015). Novak (2009) frames social justice as the promotion of collaboration and the fair allocation of benefits and burdens. From an organizational perspective, justice encompasses distributive, procedural, and interactional elements (Di Battista et al., 2014; Chory et al., 2022; Resh & Sabbagh, 2016; Rasooli et al., 2019; Thibaut & Walker, 1975).

Consequently, this study investigated the perceptions of social justice held by EFL teachers and learners within the NUEEFL context. It aims to guide the Ministry of Education, educational institutions, and test developers in improving the equity, fairness, and social responsibility of

English language assessment in Iran. As far as the researchers of the present study reviewed the literature, they could not find a similar study; therefore, this study aimed to fill the gap in the literature. On the other hand, it is necessary to explore the main English as a foreign language (EFL) stakeholders' perceptions (EFL teachers and learners) toward the language education (Montaseri et al., 2023; Ghorbani et al., 2024). Therefore, the present study explored the perceptions of EFL teachers and learners toward the social justice of the English language Konkour. As a result, the present study aimed to answer the following research questions:

1. Does the social justice questionnaire enjoy reliability and validity?
2. To what extent do EFL teachers and students perceive Konkour as a socially just test?

3. Methodology

This study adopted an explanatory sequential mixed-methods. Mixed-methods designs can provide more comprehensive approaches, in that both qualitative and quantitative methods can be used, providing a more strongly formed data set. The questionnaire-based quantitative side of the research was conducted using a questionnaire for the purposes of fairness and social justice. In addition, this research study had a qualitative phase in which the gathered data were interpreted through interviews with the test takers and the teachers. The qualitative phase of the research was reported through responses of the participants regarding the social justice of NUEEFL.

3.1. Participants

3.1.1. Quantitative Phase

3.1.1.1. B.A. Konkour Candidates

A total of 200 students (100 male and 100 female) who had taken the Iranian National University Entrance Examination for Foreign Languages (NUEEFL, or "Konkour") participated in this study. They were randomly selected from university students across different provinces in Iran, all studying at the Islamic Azad University, Science and Research Branch in Tehran. The participants were native Persian speakers, aged between 18 and 20 years old. Written consent was obtained from all students prior to their participation.

3.1.1.2. English Language Teachers

The study also included 100 English teachers with varying levels of experience and academic qualifications. Their teaching experience ranged

from 4 to 20 years, and their ages spanned from 27 to 47 years old. The teachers held different degrees, including M.A., Ph.D. candidates, and Ph.D. holders.

3.1.2. Qualitative Phase

Additionally, for the qualitative phase of the study, **30 teachers (15 male and 15 female)** and 50 students (25 male and 25 female) participated in semi-structured interviews to provide deeper insights into their perceptions of social justice in the Konkour exam.

3.2. Instruments

3.2.1. Quantitative Phase

3.2.1.1. National University Entrance Exam for Foreign Languages (NUEEFL)

The data for this research were derived from the administration of the NUEEFL in 2021, which consisted of 70 multiple-choice questions, each yielding dichotomous scores. The items were organized into six subtests, including Grammar (10 items), Vocabulary (15 items), Sentence structure (5 items), Language functions (10 items), Cloze test (10 items), and Reading comprehension (20 items).

3.2.1.2. Social Justice Questionnaire

The Social Justice Questionnaire (SJQ) focused on the extent to which the Konkour observes social justice considerations in measuring the candidates' English language ability, was developed and used. This questionnaire consisted of 14 closed-ended items with a 5-point Likert scale. The SJQ measured three constructs, including treatment of social class (items 1-6), personal bias (items 7-13), and test developer characteristics (item 14). The results of the validation of the questionnaire indicated that two factors were selected as the questionnaire's underlying constructs. The factor loadings for the 14 questions of the social justice questionnaire across the two extracted factors were collected. The first factor included items 7 to 13, which measured "personal bias". The reliability of the first factor was .912, indicating that the first factor enjoyed a high level of reliability. The results also showed that the convergent validity index for the first factor was .772. It should be noted that item 14 also loaded under the first factor, while it was supposed to measure the "test developer characteristic" factor. In addition, items 1 to 6 loaded under the second factor, which can be labeled as "treatment of social class". All factor loadings enjoyed large effect sizes. The reliability

of construct for the second factor was .908, and it also enjoyed a convergent validity of .787.

3.2.2. Qualitative Phase

3.2.2.1. Semi-structured Interview

After administering the questionnaires, the researchers of the study conducted semi-structured individual interviews with 30 teachers (15 males and 15 females) and 50 students (25 males and 25 females). The primary objective of the interview was to collect more comprehensive data about teachers' perceptions of the role of fairness, social justice, and test usefulness within the Iranian EFL setting.

3.3. Data Collection Procedure

3.3.1. Quantitative Phase

The participants were initially informed to complete the questionnaires. Following the pilot phase, the results were subjected to statistical analysis to ascertain the suitability of all test items for the actual data collection process. Each item was evaluated for validity and reliability to confirm that it effectively measured its intended constructs.

After the piloting and item analysis section, the actual data collection began. In this section, the participants were given the piloted questionnaire, and they were asked about their honest perceptions regarding the NUEEFL. It is also noteworthy that the participants were assured of the confidentiality of their responses, and they were informed that they could leave out the name section of the questionnaire blank if they wished. In addition, they were given as much time as needed to complete the questionnaire comfortably.

A mock Konkour was presented to the students to evaluate the DIF of the Konkour exam. The objective of this mock Konkour exam was to assess the test regarding DIF.

3.3.2. Qualitative Phase

After completing the questionnaire, the participants were asked about their willingness to engage in follow-up interviews. Consequently, 30 teachers (15 males and 15 females) and 50 students (25 males and 25 females) were selected for this phase of the study. The interviews were administered in person and done in Persian. Each session lasted around 20 minutes, and the interviews occurred over a span of two weeks. The researchers created the interview questions, including 10 questions, based on the questionnaire findings and pertinent literature. Subsequently, five experts in the subject, all TEFL Ph.D. holders and faculty members from

various Iranian institutions, evaluated and validated the interview questions to ensure their content validity. The interviews were transcribed after the recording, with the participants' cooperation.

3.4. Data Analysis Procedure

3.4.1. Quantitative Phase

The reliability of the SJQ was measured using the Cronbach's alpha test. Moreover, the questionnaire's construct validity was examined by an exploratory factor analysis (EFA) with the main axis factoring approach and Varimax rotation.

3.4.2. Qualitative Phase

To identify the pertinent codes and categories, the qualitative data were analyzed using thematic analysis (TA) (Braun & Clarke, 2006). TA is a method for methodically identifying, categorizing, and revealing recurring patterns of meaning (themes) in a dataset. The transcription of interview recordings was required before the coding procedures.

4. Results

4.1. Reliability and Construct Validity of Social Justice Questionnaire

The SJQ had 14 items which measured three constructs, including treatment of social class (items 1-6), personal bias (items 7-13), and test developer characteristics (item 14). The last item was designed to elicit demographic information, which is reported later.

Table 2. *The Reliability of Social Justice Questionnaire and its Components*

	Cronbach's Alpha	N of Items
Treatment of Social Class	.885	6
Personal Bias	.910	7
Total Social Justice	.859	14

Table 2 presents the reliability indicators for the SJQ and its constituent parts. The whole SJQ had a Cronbach's alpha reliability of .859. The reliability values for the treatment of social class and personal bias were .885 and .910, respectively. Based on the criteria discussed earlier, these reliability indices can be classified as "good" ($\geq .80$) or "excellent" ($\geq .90$) (George & Mallery, 2019).

The questionnaire's construct validity was examined by an exploratory factor analysis (EFA) with the main axis factoring approach and Varimax rotation.

Table 3. *Component Correlation Matrix Social Justice Questionnaire*

Component	1	2
1	1.000	
2	.122	1.000

The Varimax rotation method was employed after consulting the results of the Component Correlation Matrix. Ignoring the 1 one diagonal, the correlation between the two extracted factors was lower than $\pm.32$ (Grande, 2016; Dagdag et. al., 2020). That was why the Varimax rotation method was selected.

Figure 1. *Optimum Number of Factors for Social Justice Questionnaire*

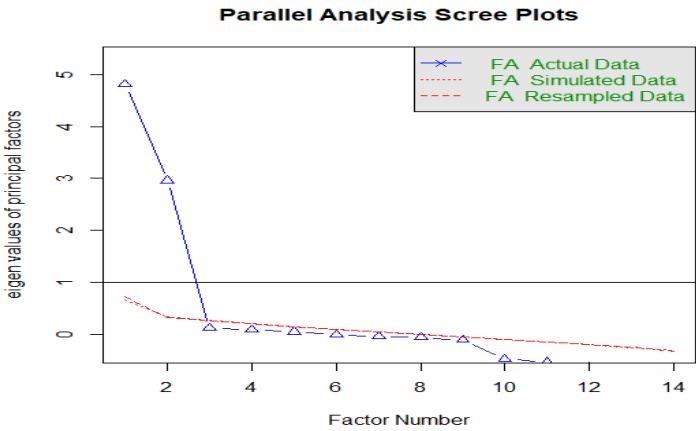


Figure 1 illustrates the optimum number of components to be retrieved via the Parallel Analysis approach (Revelle, 2020). Figure 1 illustrates that the Parallel Analysis indicated the extraction of two components as the underlying constructs of the 14 questions in the questionnaire.

Then Watkins' Parallel Analysis was conducted, which suggested two factors to be extracted as the underlying constructs of the social justice questionnaire. The original eigenvalues extracted by SPSS are available upon request. Before discussing the results of EFA, the assumptions of sampling adequacy and sphericity were checked. The KMO index of .917 was higher than the minimum acceptable criterion of .60 (Pallant, 2016; Field, 2018; Denis, 2021; Harrison et. al., 2020); hence, the adequacy of the present sample size of 266. The significant results of the sphericity test ($\chi^2(91) = 2192.62, p < .05$) indicated that the present data were factorable; i.e., there were no zero correlations among all variables. It is then obtained the number of factors that were extracted and the total variance explained

by those factors. The results indicated that two factors were selected as the questionnaire's underlying constructs. This two-factor model accounted for 59.95 percent of the total variance.

Lastly, the factor loadings for the 14 questions of the social justice questionnaire across the two extracted factors were collected. The first factor included items 7 to 13, which measured "personal bias". All factor loadings enjoyed a large effect size; that is to say, they were higher than .50 (Field, 2018). The composite reliability of the first factor was .912, indicating that the first factor enjoyed a high level of reliability. The results also showed that the convergent validity index for the first factor was .772. That is to say, there was a 77.2 percent chance that the first factor measured what it was supposed to measure. It should be noted that item 14 also loaded under the first factor, while it was supposed to measure the "test developer characteristic" factor. In addition, items 1 to 6 loaded under the second factor, which can be labeled as "treatment of social class". All factor loadings enjoyed large effect sizes. The reliability of the construct for the second factor was .908. It also enjoyed a convergent validity of .787. That is to say, there was a 78.7 percent chance that the second factor measured what it was supposed to measure.

4.2. Learners and Teachers' perceptions of the social justice of the NUEEFL

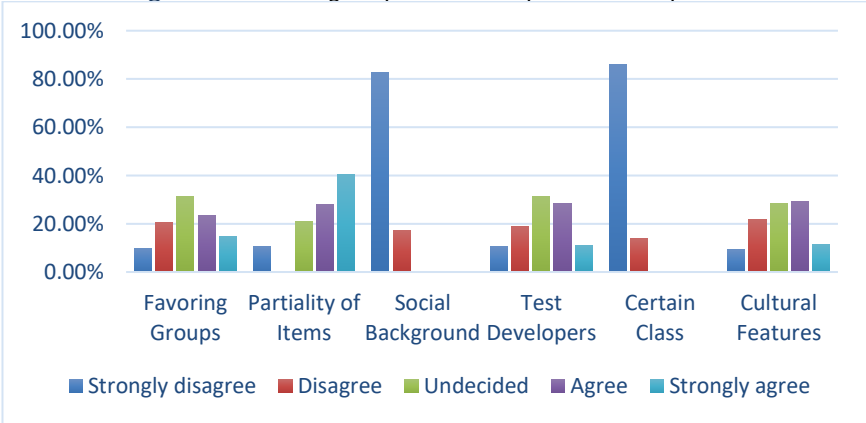
Except for the last item, which was an open-ended one, the social justice questionnaire included 14 items measuring three constructs: treatment of social class (items 1 to 6), personal bias (items 7 to 13), and test developer characteristic (item 14). The results are discussed below. Table 4 shows the frequencies and percentages for the six items measuring treatment of social class. The results indicated that 38.2 percent of respondents agreed or strongly agreed with the idea that, "the test treats a certain social class better than the others". On the other hand; 30.5 percent disagreed with the partiality of the test, and another 31.3 percent were undecided. The results also indicated that majority of the respondents; i.e., 68.5 percent, agreed with the idea "that, "items are related more to the experiences of a specific social class"; while 20.9 percent were undecided, and another 10.6 percent disagreed with the idea that items were related to some groups. All respondents disagreed or strongly disagreed with the idea that, "people with different social backgrounds may respond differently to the items of the test".

Table 4. *Frequencies and Percentages of Treatment of Social Class*

		Fairness					Total
		Strongly disagree	Disagree	Undecided	Agree	Strongly agree	
Favoring Groups	Count	26	54	82	61	39	262
	%	9.9%	20.6%	31.3%	23.3%	14.9%	100.0%
Partiality of Items	Count	28	0	55	74	106	263
	%	10.6%	0.0%	20.9%	28.1%	40.3%	100.0%
Social Background	Count	195	41	0	0	0	236
	%	82.6%	17.4%	0.0%	0.0%	0.0%	100.0%
Test Developers	Count	28	49	82	74	29	262
	%	10.7%	18.7%	31.3%	28.2%	11.1%	100.0%
Certain Class	Count	204	33	0	0	0	237
	%	86.1%	13.9%	0.0%	0.0%	0.0%	100.0%
Cultural Features	Count	24	57	74	76	30	261
	%	9.2%	21.8%	28.4%	29.1%	11.5%	100.0%
Total	Count	505	234	293	285	204	1521
	%	33.2%	15.4%	19.3%	18.7%	13.4%	100.0%

The majority of the respondents, i.e., 39.3, agreed with the idea that “the test developers are from a certain social background”. On the other hand, 29.4 percent disagreed with this idea, and another 31.3 percent were undecided. All respondents disagreed and strongly disagreed with the idea that “the test items are socially inclined towards a certain class of society”. And finally, 40.6 percent agreed with the idea that “features of certain cultures and/or languages are repeated in the test more often than others”. On the other hand, 31 percent disagreed, and another 28.4 percent were undecided.

Figure 2. *Percentages of Standards of Treatment of Social Class*



The overall results indicated that 48.6 percent of respondents disagreed with the idea that EUEE met the standards of “treatment of social class”. On the other hand, 32.1 percent held the idea that EUEE met the standards of “treatment of social class”, and another 19.3 percent were neutral. Figure 2 shows the percentages discussed above.

Table 5. *Frequencies and Percentages of Personal Bias*

		Fairness					Total
		Strongly disagree	Disagree	Undecided	Agree	Strongly agree	
Test Addressee	Count	21	0	63	84	95	263
	%	8.0%	0.0%	24.0%	31.9%	36.1%	100.0%
Neglecting Personal Qualities	Count	0	18	66	0	179	263
	%	0.0%	6.8%	25.1%	0.0%	68.1%	100.0%
Other Criteria	Count	0	24	60	0	177	261
	%	0.0%	9.2%	23.0%	0.0%	67.8%	100.0%
Subjective Results	Count	22	69	72	61	39	263
	%	8.4%	26.2%	27.4%	23.2%	14.8%	100.0%
Not Mere Performance	Count	75	84	98	4	0	261
	%	28.7%	32.2%	37.5%	1.5%	0.0%	100.0%
Unfair Treatment of Rights	Count	0	24	51	0	186	261
	%	0.0%	9.2%	19.5%	0.0%	71.3%	100.0%
Offending Content	Count	20	0	63	75	102	260
	%	7.7%	0.0%	24.2%	28.8%	39.2%	100.0%
Total	Count	138	219	473	224	778	1832
	%	7.5%	12.0%	25.8%	12.2%	42.5%	100.0%

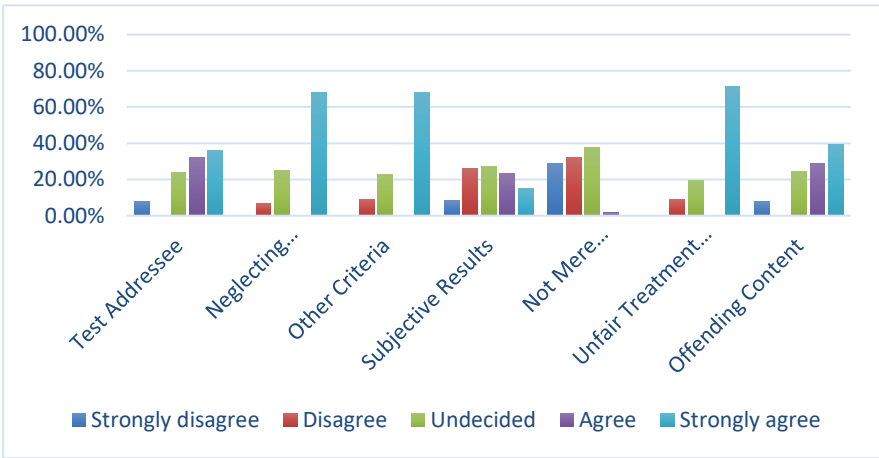
Table 5 shows the frequencies and percentages for the seven items related to personal bias. The results indicated that 68 percent believed that “I feel that the test content is not addressing me, but certain people”. On the other hand, eight percent disagreed with the idea that they were not addressees of the test, and another 24 percent were undecided.

The results also showed that 68.1 percent agreed with the idea that “my personal qualities are being neglected in the test”; 25.1 percent were undecided, while 6.8 percent held the opposite view. Most respondents (67.8 %) agreed and strongly agreed with the idea that “my performance in the test is not the only criterion for my final results”; moreover, 23 percent were undecided, and another 9.2 percent disagreed with this idea. Almost equal percentages of respondents showed opposing attitudes towards, “my final results are going to be more based on subjective than objective judgment”; while 38 percent agreed with the idea that final

results were subjective, 34.6 percent held the opposite view, and another 27.4 percent were undecided.

Most respondents, i.e., 60.9 percent, disagreed with the idea that “my mere performance is not the basis for my final results”. Only 1.6 percent held the opposite view, and 37.5 percent were neutral. The results also showed that 71.3 percent agreed and strongly agreed with the idea that “my basic rights are being treated unfairly in the test”, while 9.2 percent held the opposite opinion, and another 19.5 percent were undecided. And finally, 68 percent of respondents agreed with the idea that “Sometimes I feel offended by some of the test contents”, while 7.7 percent disagreed, and another 24.2 percent were undecided.

Figure 3. Percentages of Standards of Personal Bias



The overall results indicated that 54.7 percent of respondents agreed with the idea that EUEE met the standards of “personal bias”. On the other hand, 19.5 percent held the idea that EUEE met the standards of “personal bias”, and another 25.8 percent were neutral. Figure 3 indicates the above-mentioned percentages.

Table 6. Frequencies and Percentages of Test Developers’ Characteristics

		Fairness					Total
		Strongly disagree	Disagree	Undecided	Agree	Strongly agree	
Item 14	Count	200	34	0	0	0	234
	%	85.5%	14.5%	0.0%	0.0%	0.0%	100.0%

The last standard of social justice questionnaire, i.e., test developers’ characteristics, included a single item. As displayed in Table 6, all

respondents disagreed and strongly disagreed with the idea that, "people with certain impairments will not have an equal chance of performing in the test".

Table 7. *Areas not Treated Fairly*

	Frequency	Percent
Age	7	2.6
Gender	28	10.5
Social status	46	17.3
Ethnicity/race	29	10.9
Age, Social status	15	5.6
Gender, Social status	7	2.6
Gender, Ethnicity/race	7	2.6
Gender, Economical status	15	5.6
Social status, Ethnicity/race	7	2.6
Social status, Economical status	46	17.3
Gender, Social status, Economical status	21	7.9
Gender, Age, Social status, Economical status	16	6.0
Social status, Economical status, Ethnicity/race	22	8.3
Total	266	100.0

The last item on the social justice questionnaire was an open-ended item; i.e., "I do not know equity in the test in the following areas". Table 7 displays the frequencies and percentages of the areas the respondents recognized as being unfairly treated. The results indicated that social status (17.3 %) and social status and economical status (17.3 %) were selected by the majority of the respondents as being not treated equally in the EUEE. These were followed by ethnicity/race (10.9 %) and gender (10.5 %). The least frequent areas not treated fairly were: age, gender, and social status, gender and ethnicity/race, and social status and ethnicity/race (2.6 %).

4.3. The Interview Results

4.3.1. The Results of Teachers' Interviews

Theme 1: Socioeconomic Privilege and Structural Bias

A common idea among teachers was that the Konkour naturally favors students from richer and more privileged backgrounds. Many people said that rich students have an unfair advantage because they can afford private tutoring, international test prep programs, and a lot of extracurricular activities.

"The system helps the wealthy." "Social justice is just a slogan when a student can buy success through expensive prep courses." (T17, male)

Another teacher went into more detail:

"Parents who can afford private teachers, better schools, and ways to help their kids do well on tests are buying their kids' future." In the meantime, students from families with less money are on their own.

Participants stressed that this economic gap goes against the idea of equal opportunity. They said that the way the Konkour is set up, along with the fact that some people have more access to preparatory resources than others, keeps social hierarchies in place instead of breaking them down.

Theme 2: Language and Ideological Gatekeeping

A lot of teachers thought that using English as a way to filter students in the Konkour was a way to keep out people with different ideas. They said that requiring a high level of English proficiency without making sure that all students in the country have access to good English education unfairly hurts students in areas with less money.

"Using English to weed out students without giving them equal access to education is wrong." "It's a form of language discrimination." (T01, female)

Another teacher said:

"We say we care about fairness, but we use a foreign language as a barrier." A lot of students never got good English lessons, but we expect them to do the same work.

This theme underscores the belief that English operates as more than a mere academic discipline; it acts as an instrument for social stratification. Teachers were worried that the test favors students from elite, often urban, schools without saying so directly.

Theme 3: The Need for a Complete Change

Several teachers called for a complete overhaul of the assessment system, saying that multiple-choice tests don't fully show what students can do. They suggested adding different types of tests that could better show each person's skills, learning history, and ability to communicate.

"We need more than just a test with multiple choices to see how well students are doing." "Maybe interviews or portfolios would make it fairer." (T25, female)

Another participant said:

"The current test only checks for passive knowledge." Speaking, writing, and using language in context are all parts of real learning. "We're not testing students; we're ranking their memory," said (T02, male).

Teachers suggested more comprehensive methods that might involve ongoing evaluation, writing samples, oral interviews, and evaluation of classroom performance. They thought that these changes would not only

make the evaluation process fairer, but also lead to better teaching and learning.

In summary, teachers' views on social justice in the context of the Konkour showed that they were very worried about systemic inequalities, language barriers, and limited assessment tools. Their viewpoints indicate a yearning for systemic reform and the establishment of testing policies that are more equitable, inclusive, and pedagogically significant.

4.3.2. The Results of Students' Interviews

Theme 1: Differences in Wealth and Location

Several students noted they felt very unfair because educational resources were not evenly spread out across different regions and socioeconomic backgrounds. One male student from Tehran, for example, said, *"My cousin in Tehran had a private tutor from Canada. I didn't have anyone to even check my grammar. This is not equal opportunity."*

This shows how different it is for students to get good instruction. In big cities, students have access to international tutors and well-equipped schools, while in other places, students have to make do with little help.

Another girl from Lorestan said, *"We didn't even have a real English teacher in high school."* *"How can I compete?"*

These stories show how geography can make it harder for everyone to have the same chances, which puts rural or less wealthy students at a big disadvantage. These testimonies indicate that the Konkour, by employing uniform standards without considering these disparities, unintentionally advantages individuals with superior financial resources and access to quality education.

Theme 2: Losing faith in meritocracy

Many students were doubtful of the widely held belief that doing well on the Konkour is all about hard work and talent.

A male student said he was upset because *"They say it's about effort, but it's really about money and location that takes away all the drive"*.

Students are becoming more and more disillusioned because they feel that no matter how hard they work, things like family income and where they live will determine their chances. Many people think that the exam system makes social inequalities worse instead of making things fair for everyone. This disillusionment not only diminishes their motivation but also cultivates feelings of helplessness and injustice, potentially affecting their mental health and long-term educational goals.

Theme 3: Lack of trust in the people who give the test

Students often worried that the Konkour was not being run in a fair and open way. Some students were not sure if the way the test was given really showed how hard they had worked.

One female student said, *"I studied hard, but my friend who cheated got a better rank."* *"What is fair about that?"*

This lack of trust in the system's integrity makes students less sure that the exam will give them fair results. Others had similar concerns about favoritism, corruption, and rules not being followed equally, which made people feel even more wronged. These perceptions indicate that, in addition to structural inequalities, problems with exam management significantly influence students' beliefs that the Konkour does not promote social justice and equal opportunity.

5. Discussion

The study aimed to assess whether the NUEEFL meets social justice criteria. The results showed that the Social Justice Questionnaire (SJQ) showed high reliability (Cronbach's $\alpha > .85$ for all scales) and strong construct validity, with two main factors, personal bias and treatment of social class, accounting for most of the variance. In addition, the findings of the second research question indicated that both EFL teachers and students perceive the Konkour as lacking social justice. The findings revealed that nearly half of the respondents believe the test favors certain social and economic groups, with the majority indicating that test items and opportunities benefit privileged backgrounds. Furthermore, over half agree that the test is affected by personal bias, with concerns that factors beyond merit, such as subjective judgment and unfair treatment of rights, influence outcomes. Most respondents feel that test design does not ensure equal opportunity for all, particularly for those with impairments. The qualitative findings indicated that the teachers and students highlight systemic advantages for wealthier and urban students, language barriers, and distrust in the test's fairness and transparency. Both groups emphasize that socioeconomic status, location, and access to resources significantly affect performance, rather than ability alone.

This study is grounded in Kunnan's Test Fairness Framework (TFF), which outlines that fairness in testing comprises five interconnected dimensions: validity, absence of bias, access, administration, and social consequences (Kunnan, 2004, 2010). In the present investigation, although SJQ exhibits considerable validity, this psychometric robustness does not carry over to the Konkour examination. Initially, concerning the lack of bias, both EFL educators and learners recognize that the Konkour tends to favor those from socioeconomically advantaged backgrounds and

is shaped by subjective evaluations and inequitable treatment of rights, indicating the existence of construct-irrelevant bias that compromises fairness. Secondly, regarding access, the assessment seems to place examinees from lower socioeconomic backgrounds, rural areas, or those with limited resources at a disadvantage, suggesting unequal opportunities for preparation and participation in the exam. Third, while the quantitative data did not directly evaluate administration, qualitative narratives indicate a pervasive skepticism regarding the fairness and transparency of the test. This situation prompts significant concerns about procedural equity and the consistency of testing conditions. Ultimately, regarding social consequences, the assessment appears to perpetuate current social hierarchies, privileging urban and wealthy students while challenging the ideals of meritocracy, which resonates with Kunnan's focus on the wider societal effects of testing methodologies.

The SJQ exhibited strong reliability and construct validity. The Cronbach's alpha values, which span from .859 to .910 for both the total scale and its subscales, indicate that the SJQ adheres to recognized benchmarks for strong to exceptional internal consistency (George & Mallery, 2019). The exploratory factor analysis revealed a distinct two-factor structure, encompassing personal bias and the treatment of social class, which together accounted for almost 60% of the total variance, with robust factor loadings (all exceeding .50). This psychometric robustness corresponds with recent validation studies in language testing, highlighting the significance of strong reliability and construct validity in tools that evaluate fairness and equity (Khan & Javed, 2025).

The perceptions held by teachers and students regarding the Konkour highlight significant issues related to social justice and equity. Quantitative findings reveal that almost half of the respondents perceive the test as favoring specific social classes, whereas a majority recognize the impact of personal biases on the results. Participants also pointed out disparities associated with disabilities, economic status, and ethnic background. The findings align with existing literature regarding fairness in EFL assessment, emphasizing the importance of equal access, transparency, and the reduction of demographic biases in shaping perceptions of justice (Tofighi & Ahmadi Safa, 2023; Rasooli et al., 2022).

In-depth qualitative insights enhance comprehension of these perceptions. The educators highlighted that the Konkour intrinsically benefits wealthy urban students who have access to private tutoring and resources, thereby perpetuating systemic inequalities. The students articulated this sentiment, voicing their frustrations regarding the

disparities in educational opportunities and a diminishing belief in the principles of meritocracy. The narratives correspond with the findings of Behboudi Nazhame et al. (2024), who also recognized validity, fairness, and structural barriers as significant concerns in the General English University Entrance Examination in Iran.

6. Conclusion

The researchers of the current study concluded that the SJQ is a reliable and valid instrument, exhibiting strong internal consistency and solid construct validity. This study's empirical findings offer fresh insights into the perceptions of EFL teachers and students regarding the Konkour exam. Both groups viewed it as socially unjust, emphasizing significant inequities related to socioeconomic status, geographic location, and language barriers. This study has illuminated the ways in which structural disparities are evident in high-stakes testing. The participants highlighted that privileged students gain advantages from superior resources, while others encounter systemic obstacles. This reinforces the notion that the Konkour undermines meritocratic principles and does not produce equitable results.

The results can have some implications for teachers, test developers, and the mainstream education system, especially the Ministry of Education of Iran. These implications are described here in this section one by one. One of the implications that can be made from the results is for EFL teachers. By studying the findings, EFL teachers can become aware of the factors that have an impact on their students' performance in the Konkour, which may eventually lead to their future, especially finding a suitable job. Becoming aware of the shortcomings of the test and the factors that lead to some bias can be a very important factor for improving it by adjusting the expectations towards the test. The results can help the Ministry of Education and the Ministry of Science, Research and Technology to make a better selection of students so that students can be sure that they are treated fairly in the test that leads to their academic performance and finding a suitable job in the related fields. As a result, if the government becomes aware of the problems that exist in the test, they can try to increase the performance and justice, and fairness of the tests by considering the factors under study in this research. Test developers are the other beneficiaries of this study. By studying the results of this study, test developers can find out the disadvantages and shortcomings of the Konkour and see how they can improve their tests and items they develop. Therefore, if they know what factors can have an impact on examinees' fair treatment based on social justice, usefulness, and fairness

in terms of different social factors, they can try to eliminate such problems.

Future studies could rigorously examine the concept of fairness in high-stakes EFL assessment by involving a variety of stakeholders and utilizing participatory methods, such as Appreciative Inquiry, to collaboratively develop equitable testing practices. In addition, researchers might explore various assessment formats, like portfolios, oral tasks, and coursework, to mitigate systemic bias, while also examining how contextual elements (e.g., teacher background, geographic location) shape perceptions of fairness. Furthermore, the creation and validation of assessment frameworks focused on social justice, specifically designed for contexts such as the Konkour, would enhance both theoretical understanding and practical application in the realm of equitable testing.

Funding: This research received no external funding from any agency.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Ahmadi Safa, M., & Beheshti, S. (2025). Revised test fairness framework (RTFF): Modeling and structure. *Language Testing in Asia*, 15(54), 1–20. <https://doi.org/10.1186/s40468-025-00393-6>
- Beheshti, S., & Ahmadi Safa, M. (2023). Reconceptualization of test fairness model: A grounded theory approach. *Iranian Journal of Language Teaching Research*, 11(2), 119–146. <https://doi.org/10.30466/ijltr.2023.121333>
- Blake, C. (2015). *Teaching social justice in theory and practice*. Concordia University–Portland. <https://education.cu-portland.edu/blog/classroom-resources/teaching-social-justice/>
- Camilli, G. (2006). Test fairness. *Educational Measurement*, 4, 221–256.
- Camilli, G. (2018). Ongoing issues in test fairness. *Educational Research and Evaluation*, 19(2–3), 104–120. <https://doi.org/10.1080/13803611.2013.767602>
- Chan, D., Jennings, D., Clause, C. S., & Delbridge, K. (1998). Applicant perceptions of test fairness: Integrating justice and self-serving bias perspectives. *International Journal of Selection and Assessment*, 6(4), 232–239. <https://doi.org/10.1111/1468-2389.00094>
- Dagdag, J. D., Anoling Jr, O. C., Salviejo, R. P., Pascual, J. F., & Dagdag, J. M. H. (2020). Development of problem-solving efficacy scales in mathematics. *Universal Journal of Educational Research*, 8(6), 2397–2405. <https://doi.org/10.13189/ujer.2020.080624>

- Davies, A. (2010). Test fairness: A response. *Language Testing*, 27(2), 171–176. <https://doi.org/10.1177/0265532209349466>
- Denis, D. J. (2021). *Applied univariate, bivariate, and multivariate statistics: Understanding statistics for social and natural scientists, with applications in SPSS and R*. John Wiley & Sons.
- Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.
- George, D., & Mallery, P. (2019). *IBM SPSS statistics 26 step by step: A simple guide and reference*. Routledge.
- Ghorbani, S., Afraz, S., Fatehi Rad, N., & Samimi, F. (2024). Exploring Iranian EFL learners' perceptions of critical pedagogy-based teacher effectiveness. *Journal of Language, Culture, and Translation*, 5(2), 218–240. <https://doi.org/10.71864/LCT-2024-1183176>
- Grande, T. L. (2016). *Selecting a rotation in a factor analysis using SPSS* [Video]. YouTube. <https://youtube.com>
- Harrison, V., Kemp, R., Brace, N., & Snelgar, R. (2020). *SPSS for psychologists*. Bloomsbury Publishing.
- Hossain, M. M. (2018). Implementing social justice in teaching and learning English at university level in Bangladesh: Necessities, approaches and challenges. *International Journal of English Language and Literature Studies*, 7(1), 7–14. <https://doi.org/10.18488/journal.23.2018.71.7.14>
- Isbell, D. R., Kremmel, B., & Kim, J. (2023). Remote proctoring in language testing: Implications for fairness and justice. *Language Assessment Quarterly*, 20(4–5), 469–487. <https://doi.org/10.1080/15434303.2023.2288251>
- Kamyab, S. (2008). The university entrance exam crisis in Iran. *International Higher Education*, (51). <https://doi.org/10.6017/ihe.2008.51.8010>
- Kamyab, S. (2007). Flying brains: A challenge facing Iran today. *International Higher Education*. <https://doi.org/10.6017/ihe.2007.47.7955>
- Karataş, T. O., & Okan, Z. (2021). A conceptual framework on the power of language tests as social practice. In B. Lantaigne, C. Coombe, & J. D. Brown (Eds.), *Challenges in language testing around the world* (pp. 47–56). Springer Nature.
- Kunnan, A. J. (2004). Test fairness. In M. Milanovic & C. Weir (Eds.), *European language testing in a global context* (pp. 27–48). Cambridge University Press.
- Kunnan, A. J. (2005). Towards a model of test evaluation: Using the test fairness and test context frameworks. In *Multilingualism and*

- assessment: *Achieving transparency, assuring quality, sustaining diversity. Proceedings of the ALTE Berlin Conference* (pp. 229–251).
- Kunnan, A. J. (2010). Test fairness and Toulmin's argument structure. *Language Testing*, 27(2), 183–189. <https://doi.org/10.1177/0265532209349468>
- Matthew, P. W., & Coral, Y. Q. (2021). Language classroom assessment fairness: Perceptions from students. *LEARN Journal: Language Education and Acquisition Research Network*, 14(1), 492–521. <https://so04.tci-thaijo.org/index.php/LEARN/article/view/248700>
- Roever, C., & McNamara, T. (2006). *Language testing: The social dimension*. Blackwell Publishing.
- Montaseri, A., Pajooan Doost, R., & Ghahremani, D. (2023). An exploration into Iranian experienced EFL teachers' perceptions toward a new model of teacher supervision. *Journal of Language, Culture, and Translation*, 5(2), 190–217. <https://doi.org/10.30495/LCT.2023.1996461.1100>
- Naseriazar, A., & Badrian, A. (2010). The university scene. In *Getting into varsity: Comparability, convergence and congruence* (p. 169).
- Novak, M. (2009). Three precisions: Social justice. *First Things*. <http://www.firstthings.com/onthesquare/2009/12/three-precisions58-social-justice>
- Pallant, J. (2016). *SPSS survival manual* (6th ed.). Allen & Unwin.
- Parviz, M. (2023). Reflecting on the consequences of the Iranian university entrance examination: A systematic-narrative hybrid literature review. *Discover Education*, 2(22). <https://doi.org/10.1007/s44217-023-00046-x>
- Revelle, W. (2020). *Psych: Procedures for psychological, psychometric, and personality research* (Version 1.9.12). Northwestern University. <https://CRAN.R-project.org/package=psych>
- Rodas, C. R., & Osborn, A. (2016). Social justice and language acquisition for English language learners. In C. R. Rodas & A. Osborn (Eds.), *Social justice instruction: Empowerment on the chalkboard* (pp. 325–337).
- Sadeghi, M., Jalali, V., & Fatehi Rad, N. (2023). Appraising Iranian EFL learners' perceptions towards the teachers' implementation of classroom fairness: Focus on the main problems. *Journal of Language, Culture, and Translation*, 5(2), 175–189. <https://doi.org/10.30495/LCT.2023.1997560.1101>
- Safari, P. (2016). Reconsideration of language assessment is a MUST for democratic testing in the educational system of Iran. *Interchange*, 47(3), 267–296. <https://doi.org/10.1007/s10780-016-9276-8>

- Sharifian, F. (2012). Linguistic theory and cultural conceptualisations. *Journal of Language, Culture, and Translation*, 1(3), 93–110.
- Song, X. (2016). Fairness in educational assessment in China: Historical practices and contemporary challenges. In S. Scott, D. E. Scott, & C. F. Webber (Eds.), *Assessment in education: Implications for leadership* (pp. 67–89). Springer. https://doi.org/10.1007/978-3-319-23398-7_4
- Taheri-Larki, M., & Eidy, R. (2022). A study of demotivational factors in the Iranian ELT context: Resolutions and attitudes from teachers. *Journal of Language, Culture, and Translation*, 4(2), 184–208. <https://doi.org/10.30495/LCT.2022.699176>
- Takala, S., & Kaftandjieva, F. (2000). Test fairness: A DIF analysis of an L2 vocabulary test. *Language Testing*, 17(3), 323–340. <https://doi.org/10.1177/026553220001700303>
- Xi, X. (2010). How do we go about investigating test fairness? *Language Testing*, 27(2), 147–170. <https://doi.org/10.1177/0265532209349465>