Research Paper

Developing and Validating a Teacher Classroom Justice Scale for the Iranian EFL Context

Masoomeh Estaji[1], Kiyana Zhaleh [2], & Chiara Berti[3]

[1]Corresponding author, Department of English Language and Literature, Faculty of Persian Literature and Foreign Languages, Allameh Tabataba'i University, Tehran, Iran; *mestaji74@gmail.com*

[2]Department of English Language and Literature, Faculty of Persian Literature and Foreign Languages, Allameh Tabataba'i University, Tehran, Iran; *k_zhaleh97@atu.ac.ir*

[3]Department of Psychological, Health & Territorial Sciences, University of Chieti-Pescara, Via dei Vestini 29 - 66100 Chieti, Italy; *chiara.berti@unich.it*

Abstract

Classroom justice is the degree of perceived fairness in the distribution of outcomes, enactment of procedures, and teacher-student relationships in classrooms. This study aimed to develop and validate a Teacher Classroom Justice Scale (TCJS). After thoroughly reviewing the extant literature, scrutinizing the existing questionnaires, and interviewing experts in the field, a draft version of the instrument involving 46 items was developed and pilot-tested with 30 Iranian EFL teachers. Subsequently, another group entailing 398 Iranian EFL teachers answered the scale, and reliability was examined for each of its components. Subsequently, Exploratory Factor Analysis (EFA) revealed that a three-factor solution about procedural, interactional, and distributive justice could best explain the scale. Finally, the EFA results were approved through Confirmatory Factor Analysis (CFA), which showed that the finalized TCJS consists of 18 items and enjoys good psychometric properties of validity and reliability. Language instructors, researchers, and practitioners can use the present study findings by employing the TCJS to assess the perceptions of classroom justice in the particular domain of second/foreign language (L2) education.

Keywords: Dimensions of Classroom Justice; English as a Second/Foreign Language; Justice Principles; Scale Validation; Teachers' Justice Perceptions.

1. Introduction

In the education system, there is a heavy burden on the teachers' shoulders as they are considered its most pivotal figure, influencing both the academic accomplishments and well-being of students and the appropriate functioning of the system as a whole (Zhaleh et al., 2018). Thus, nurturing teacher professional quality is important for determining desirable processes of instruction and learning (Derakhshan et al., 2022; Mirshojaee et al., 2019). One of the characteristics directly contributing to teachers' professional effectiveness is their just practice in the instructional context (Chory et al., 2017). Teachers' just treatment of students is important because teaching is a moral undertaking whereby the responsibility of teachers is not confined to enhancing the learning of a particular subject but, more importantly, includes transmitting democratic and ethical values such as fairness and justice to students (Kazemi, 2016). Attainment of these values is significant to students, and in support of this argument, research evidence has substantiated that being treated fairly by instructors is among the main priorities of students (Dalbert, 2013; Mameli et al., 2018). Thus, studying teacher classroom justice behavior – whether at the distributive level of allocating educational resources among students, the procedural level of enacting classroom rules, procedures, and processes, or the interactional level of sharing information and transactionally communicating with learners – is of paramount concern. As mentioned by Sabbagh and Resh (2016), "education is a distinct sphere of justice" (p. 2), and for maintaining justice in this realm, education-specific justice principles need to be activated by the teachers. Nevertheless, students from primary, secondary, or even tertiary levels of

education often report experiencing teacher injustice (Čiuladienė & Račelytė, 2016; Gasser et al., 2018). Since the teachers' unjust treatment of students can have undesirable implications for the emotional, social, behavioral, and academic functioning of the students (Chory et al., 2017), teachers' attending to this crucial aspect of their professional behavior is very crucial.

The importance of teacher justice is felt substantially in Second/Foreign Language (L2) learning and instruction processes because of their social as well as relational nature (Mercer & Dörnyei, 2020). To maintain the relational cornerstone of L2 education, language teachers are expected to build positive teacher-student interpersonal relationships (Gasser et al., 2018) by meeting such qualities as mutual respect, equality, enjoyment, honesty, trust, reciprocity, warmth, open communication, and reliability (Roffey, 2011), which overlap with the principles of justice in education. However, despite its essentiality, teacher justice is an underrepresented area of research in L2 education. The first studies on classroom justice in L2 education were conducted very recently by Estaji and Zhaleh (Chory et al., 2022; Estaji & Zhaleh, 2021a, 2021b, 2022a, 2022b, 2022c; Zhaleh et al., 2022), who engaged in in-depth explorations of challenges, experiences, as well as perceptions of Iranian EFL instructors and students regarding justice and injustice in L2 classes. As Rasooli's (2020) systematic review of studies on classroom justice revealed, questionnaires and scales have been the dominant instruments utilized for measuring the perceptions of teacher justice in the instructional context. However, there is a dearth of such scales specific to the L2 education context, which may justify why empirical evidence on classroom justice is scanty in this domain.

To occupy the identified gap, regarding the need for constructing a proper teacher justice scale in the EFL context and precipitating the expansion of survey-based studies on classroom justice in L2 education, the present study sought to develop and validate a Teacher Classroom Justice Scale (TCJS) for measuring perceptions that EFL instructors in Iran have toward their practice of justice in language classes.

2. Literature Review

2.1. Social Psychology of Justice

Justice was originally studied from a socio-psychological perspective in the realms of organizational behavior, political sciences, and social sciences (Tyler, 1987). In the organizational domain, organizational justice deals with understanding one's fairness perceptions concerning the processes and outcomes happening in the workplace context (Cropanzano & Greenberg, 1997). It entails the three dimensions of distributive, interactional, and procedural justice (Di Battista et al., 2014). Distributive justice refers to the degree of perceived fairness in the received outcomes and resources, resting on the three principles of *equity* (providing resources and outcomes considering one's endeavors and performance), *equality* (sharing outcomes or resources equally among all), and *need* (sharing resources or outcomes based on persons' unique needs and exceptionalities) (Adams, 1965; Chory et al., 2022).

Procedural justice has to do with the degree of perceived fairness about the procedures and processes employed to make decisions about the distribution of outcomes (Resh & Sabbagh, 2016). This dimension rests on the principles of *bias suppression* (processes are perceived to be unbiased), *correctability* (procedures are correctable), *voice* (procedures are arrived at by considering all the individuals' opinions and concerns), *ethicality* (rules are enacted based on some ethical criteria), *transparency* (processes are performed explicitly and clearly), *accuracy* (processes are based on precise and adequate information), *reasonableness* (procedures are reasonable), and *consistency* (processes are enacted invariably across persons and time) (Rasooli et al., 2019; Thibaut & Walker, 1975).

Interactional justice pertains to degree of fairness perceived about transactional and informational communication among persons (Chory, 2007). Interactional justice entails the principles of *respect* (respectful treatment of others), *timeliness* (on-time imparting of information), *caring* (having a caring relationship with others), *justification/adequacy* (presenting justifiable and sufficient information), *propriety* (behaving with decency), and *truthfulness* (imparting information honestly and truthfully) (Bies & Moag, 1986; Estaji & Zhaleh, 2022).

Following this conceptualization, a number of key scholars extended the line of justice research to the sphere of education (e.g., Chory-Assad & Paulsel, 2004; Dalbert & Stoeber, 2006; Sabbagh & Resh, 2016) which has resulted in a burst of studies in this area in the West over the last 20 years (e.g., Horan et al., 2010; Mameli et al., 2020; Sonnleitner & Kovacs, 2020; Tripp et al., 2019). Accordingly, the concept of classroom justice was introduced (Chory, 2007; Chory-

Assad & Paulsel, 2004) for explaining students' or instructors' fairness perceptions toward (1) distributing outcomes such as grades, rewards, feedback, and teacher time (i.e., *distributive classroom justice*), (2) enacting procedures and policies such as attendance policy, syllabus design, or grading criteria (i.e., *procedural classroom justice*), and (3) imparting information and making teacher-student interpersonal relationships (i.e., *interactional classroom justice*). It has also been posited that the three dimensions of justice and their respective principles can be enacted or violated in any classroom domain, including learning, teaching, assessment, and interactions (Estaji & Zhaleh, 2021b; Rasooli et al., 2018).

2.2. The Existing Classroom Justice Instruments

The dominant measurement instrument of classroom justice perceptions has been questionnaires (Estaji & Zhaleh, 2021a), although the existing questionnaires come with their limitations. In a systematic review of classroom justice instruments, Rasooli (2020) scrutinized 97 quantitative studies which had employed classroom justice questionnaires. His results indicated that only a small number of these studies did rigorous statistical analyses and pursued best practice procedures to support their validity interpretations. Furthermore, no domain-specific scale exists to measure classroom justice perceptions in EFL instructional contexts. Moreover, the majority of these scales (e.g., Chory-Assad & Paulsel, 2004; Ehrhardt et al., 2018; Sonnleitner & Kovacs, 2020) are locally bound and developed in the West, based on mainly the Anglo-European culture. Thus, they are not directly applicable to non-European cultures in other parts of the world such as the Middle East, Asia, or Africa.

More importantly, the extant literature on classroom justice is replete with studies assessing the students' perceptions of justice, employing student-perceived classroom justice scales of Chory (2007), Chory-Assad and Paulsel (2004), Dalbert and Stoeber (2006), Di Battista et al. (2014), and Gorard (2012), among others. Nevertheless, as both instructors and students play a crucial role in and are affected by the experience of justice and injustice, evaluating the teachers' perceptions is as important as examining those of the students. As stated by Derakhshan et al. (2020), teachers are the most influential figures in the students' academic lives, and they are the main agents responsible for ensuring justice in the classroom (Sabbagh & Resh, 2014). Thus, for facilitating the expansion of studies on the teachers' perceptions of justice, there is a desideratum for developing a well-designed and valid teacher-perceived justice scale in different fields of studies in both Western and non-Western educational contexts.

As stated by Cropanzano et al. (2015), to reach a more accurate judgment of a situation as just or unjust, one should examine whether and to what extent the three dimensions of justice are realized through the application of their justice principles (e.g., need, quality, voice, bias suppression, adequacy principle). Nevertheless, the two most frequently used classroom justice scales, developed by Chory (2007) and Chory-Assad and Paulsel (2004), are direct fairness measures in general, assessing interactional, procedural, and distributive dimensions as perceived by students without referring to the justice principles. Ehrhardt-Madapathi et al. (2018) asserted that this mere focus on the perceptions of fairness to the disregard of unique justice principles accounts for a limited and subjective conceptualization of classroom justice. To address this concern, researchers have developed some scales (e.g., Berti et al., 2010; Di Battista et al., 2014; Gorard, 2012; Kazemi, 2016) measuring procedural, distributive, or interactional justice through the justice principles. However, no comprehensive scale has been developed to date, measuring the three-justice dimensions and all their relevant principles in all classroom domains.

As a stride toward addressing the identified lacunas and limitations, the present study aimed to develop and validate a teacher-perceived classroom justice scale in a particular non-Western educational context (i.e., the Iranian EFL context), catering for the three classroom justice dimensions and their respective principles.

3. Purpose of the Study

Based on the importance of incorporating justice in language classes and lack of a valid and comprehensive scale for assessing how EFL teachers perceive their own classroom justice with regard to the enactment of justice dimensions and all their principles in the learning, teaching, assessment, and interaction domains, this study intended to develop and validate a teacher classroom justice scale in the Iranian EFL context. More particularly, the researchers formulated the following research question in this study:

Does the Teacher Classroom Justice Scale (TCJS) demonstrate the psychometric properties of validity and reliability?

Shahid Chamran University of Ahvaz

4. Methodology

4.1. Participants

In the current study, the participants, chosen through convenience sampling, were 398 Iranian EFL instructors of private language institutions in various provinces of Iran. They were from both genders and differed with regard to their academic degrees, major, age, teaching experience, and teaching levels. The rationale for targeting the participants from various levels, groups, and locations was to facilitate generalizing the current study results. The participants' information is presented in Table 1.

Table 1. *The Participants' Demographic Information*

Demographic information	Participants (f)
Gender	
Female	247
Male	151
Age	
Less than 20	8
20-29	185
30-39	127
40-49	70
50 or more	8
Academic Degree	
High School Diploma	30
Associate of Arts	11
BA	127
MA	187
PhD	43
Years of Teaching Experience	
0-4	138
5-9	105
10-14	59
15-19	35
20-24	30
25 or more	31
Teaching Levels	
Beginner	161
Early intermediate	203
Intermediate	254
Advanced	161
Proficient	66
Children	122
Teenagers	286
Adults	206
Majors	
Teaching English as a Foreign Language	317
English Language and Literature	39
English Translation	28
Linguistics	7
Majors Other Than English	7

4.2. Instrumentation

Teacher Classroom Justice Scale (TCJS). To carry out the study, a TCJS was primarily developed by the researchers and subsequently validated by collecting data from 398 EFL teachers. This scale had two parts. In the first part, the participants' demographic information was gathered (see Appendix A). In the second part, the constructed questionnaire entailed 46 items to be answered on a range of 1 (Never) to 5 (Always). The procedure and steps taken for the questionnaire development and validation are presented in the following section.

4.3. Procedure

To carry out the study, at first, TCJS was developed following the three standard procedures for developing a measurement instrument with good psychometric properties (Dörnyei, 2003; Ghaedsharafi et al., 2019); namely: (1) reviewing the classroom justice literature, (2) scrutinizing the existing questionnaires on classroom justice, and (3) interviewing experts in the field.

Classroom justice comprises the three main dimensions of interactional, procedural, and distributive justice, each being realized through some unique principles. Hence, distributive justice was realized through the equality, need, and equity principles (Deutsch, 1975). Interactional justice was actualized through caring, propriety, respect, justification/adequacy, timeliness, and truthfulness principles (Bies & Moag, 1986; Rasooli et al., 2019). Procedural justice was enacted through bias suppression, ethicality, correctability, accuracy, voice, transparency, consistency, and reasonableness principles (Rasooli et al., 2019; Thibaut & Walker, 1975). Finally, it was found that justice dimensions and their respective principles apply in the learning, interactions, teaching, and assessment domains of the classroom (Chory et al., 2017; Horan et al., 2010; Rasooli et al., 2019). Accordingly, a three-level teacher classroom justice framework was developed by Estaji and Zhaleh (2021a), about the dimensions, principles, and domains of classroom justice. This framework was adopted for developing the TCJS in the present study.

Having reviewed the literature, the present study researchers held one-on-one, semi-structured interviews with five experienced EFL teachers, who were considered experts in teaching and were engaged in educational research undertakings as a part of their professional practice. The rationale behind holding interviews was to examine if the interviewees approved of the elements found important in the developed classroom justice framework (Estaji & Zhaleh, 2021a), if they could add other significant elements to the framework, and construct some items for the TCJS uniquely based on the interviewees' accounts and perceptions regarding language teachers' classroom justice behaviors. The interview prompts pertained to the chief constructs of the scale, attending to the various dimensions of classroom justice (Appendix B). To ensure the trustworthiness of the prompts (Nassaji, 2020), three university professors experienced in doing educational research examined them regarding language clarity and content relevance criteria. After applying the experts' feedback, the researchers modified the prompts, and their content validity was ensured. Each interview session lasted for about 20 or 30 minutes. The sessions were recorded for later transcription. The interview sessions were run online through Skype or WhatsApp due to face-to-face inaccessibility to participants because of the Covid-19 outbreak. In compliance with Gao and Zhang (2020), interview data was analyzed by following the step-by-step procedures of cleaning and coding the data, generating and categorizing themes, and finally reporting and interpreting the obtained codes and themes. It was found that the framework developed by Estaji and Zhaleh (2021a) and the themes and codes obtained from the content analysis of the interviews were in line with each other and assisted the researchers in the development of some items for the scale.

Subsequently, the existing scales on classroom justice (Berti et al., 2010; Chory, 2007; Chory-Assad & Paulsel, 2004; Dalbert & Stoeber, 2006; Di Battista et al., 2014; Ehrhardt et al., 2018; Gorard, 2012; Gouveia-Pereira et al., 2003; Kazemi, 2016; Lizzio et al., 2007; Pnevmatikos & Trikkaliotis, 2012; Resh & Sabbagh, 2017; Robbins & Jeffords, 2009; Sonnleitner & Kovacs, 2020) were scrutinized to identify the relevant items for developing the questionnaire.

Based on reviewing the literature, interviewing the experts in the field, and scrutinizing the existing scales, 46 items were developed, the responses to which could range from 1 "never", 2 "rarely", 3 "often", 4 "usually", to 5 "always". Ten university professors with prior experience in developing questionnaires evaluated the 46 scale items regarding language clarity and content relevance criteria. Only a few recommendations were proposed by the experts to improve language clarity (e.g., *[item 42]: Modify "I have no favorite students that I prefer to others in class" to "I do not have a favorite student who I treat differently from other students in class"*) and content relevance of the items (e.g., *[item 39]:*

Modify "I set reasonable expectations for my students' performance based on their abilities" to "I set reasonable expectations for my students' performance"). Accordingly, the researchers modified and finalized the items based on the experts' comments, and in this way, checked their content validity.

The items were presented randomly in the questionnaire to avoid any bias in item ordering. The items corresponding to the three major factors constituting the questionnaire; namely, *Distributive Justice* (Items 1, 3, 8, 12, 13, 14, 15, 17, 22, 23, 26, 30, 42, and 46); *Procedural Justice* (Items 4, 5, 10, 11, 16, 18, 24, 32, 33, 34, 35, 36, 38, 39, 40, 41, 43, 44, and 45); and *Interactional Justice* (Items 2, 6, 7, 9, 19, 20, 21, 25, 27, 28, 29, 31, and 37). To strengthen the scale reliability, some of the items (Items 9, 14, and 37) were negatively worded, and reverse scoring was applied to them. Appendix C presents the 46 items of the questionnaire and the justice dimension, principle, and domain that each item belongs to. Clear instructions regarding the aim of the scale, the rating scale, and how to answer the items were presented at the outset of the questionnaire.

Due to the prevalence of Covid-19 at the time of data collection, to observe the health protocols and prevent the dispersion of the Corona Virus, the researchers gathered all the data virtually. In this regard, the consent letter, the demographic information scale, and the TCJS were prepared in Google Forms and sent to the potential participants (N= 398) via Telegram, email, WhatsApp, or other social media. In compliance with Helsinki's (1964) declaration about the ethics principles in conducting research studies, before filling out the scales, the participants signed a formal consent letter, indicating their voluntary participation (BERA, 2011). The researchers presented the data anonymously. No economic incentive was offered. It took participants around 15 to 20 minutes to respond to the scales.

The instruments were prepared in English since the targeted participants were EFL instructors with an adequate level of proficiency. To analyze the data, SPSS (version 24) was used to estimate the reliability and perform Exploratory Factor Analysis (EFA), and IBM AMOS (version 24) was run to perform Confirmatory Factor Analysis (CFA) on the data.

5. Results

5.1. Pilot Study

Initially, a pilot study was conducted with 30 respondents who were as similar as possible to the participants of the main phase of the study regarding their demographic information. The rationale of the pilot study was to both identify the unexpected problems, if any, in the administration of the questionnaire and reach an a priori estimation of reliability along with the identification of items that could lower reliability in the main administration. Cronbach's alpha reliability and Cronbach's alpha reliability based on standardized items for the questionnaire in the pilot study were found to be .90 and .92, respectively. The estimated reliability turn-out of .90 was a high index. In the next step, the item total statistics was inspected for items that lower the reliability. To do so, items with correlations lower than .30 to the overall reliability were identified. It was found that none of the items had negative contributions to the overall reliability. The omission of the few items that had correlations below .30 (Items 1, 3, 10, 12, 18, 20, 33, and 41) also made slight positive changes in the overall reliability. One possible reason for the low correlations of the specified items could be the low variance resulting from the small sample size in the pilot study. Therefore, the researchers concluded that all the items should be included in the main study.

5.2. Pre-Processing of the Data

The researchers started by screening the data for missing data and unengaged responses. As stated before, there were 398 responses to the TCJS. In the data, there was no missing answer. Initially, the data were checked for patterns (constant, increasing, or decreasing). Consequently, one case (case No. 311) with a constant pattern of answering was detected and removed. Subsequently, the standard deviation regarding each component was inspected from the respondents' answers, and 24 cases whose answers had low standard deviations of below .30 (Tabachnick & Fidell, 2018) were excluded as they were considered unengaged respondents (cases No. 17, 71, 114, 118, 131, 137, 142, 191, 198, 206, 214, 228, 236, 253, 307, 314, 318, 327, 337, 347, 348, 354, 355, and 377). Finally, for the main data analysis, 373 respondents remained.

5.3. Reliability Analysis

Before running factor analysis, the reliability of the scale within each component was examined. The item-total statistics for each item in the three components was inspected. As recommended by Field (2018), the item-total correlation should not be less than .30" (p. 1050). Similarly, Pallant (2016) stated, "low values (less than .3) here indicate that the item is measuring something different from the scale as a whole" (p. 119). Therefore, in the current study, those items whose item-total correlation was below .30 and whose omission did not negatively change the reliability were removed.

Table 2. *The Item-Total Statistics of Reliability for the Distributive Justice Component*

Item	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q01	54.44	39.66	.15	.14	.73
Q03	53.55	39.00	.39	.23	.71
Q08	54.21	39.78	.17	.15	.73
Q12	54.21	39.03	.27	.19	.72
Q13	54.02	37.97	.33	.18	.71
Q14	53.54	36.88	.56	.38	.69
Q15	53.70	35.71	.47	.39	.70
Q17	53.56	38.01	.38	.21	.71
Q22	54.39	38.79	.22	.13	.73
Q23	53.47	37.32	.60	.43	.69
Q26	54.00	36.45	.33	.29	.72
Q30	53.70	35.33	.43	.41	.70
Q42	54.19	37.47	.29	.20	.72
Q46	53.58	38.56	.38	.26	.71

First, the distributive component was examined, and the initially estimated reliability turned out to be .73. Subsequently, the item-total statistics was checked for each item (Table 2). According to Table 2, the omission of two items (Items 1 and 8) having low item-total correlations (i.e., below .30) could improve the reliability. Therefore, these items were removed. However, although items 12, 22, and 42 had low item-total correlations (i.e., below .30), they were not omitted as their omission changed the reliability negatively. The Cronbach's alpha reliability and Cronbach's alpha based on standardized items for the distributive justice component became .74 and .76, respectively, after removing the two items.

Next, the reliability of the procedural justice component was examined. The initial reliability for this component turned out to be .82. Subsequently, the item-total statistics was inspected (Table 3).

Table 3. *The Item-Total Statistics of Reliability for the Procedural Justice Component*

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q04	73.87	68.81	.44	.27	.81
Q05	73.57	69.16	.42	.26	.81
Q10	75.71	73.48	.02	.06	.84
Q11	74.39	70.88	.20	.05	.83
Q16	74.26	70.01	.19	.10	.83
Q18	73.49	68.06	.35	.20	.82
Q24	73.66	68.27	.44	.25	.81
Q32	73.60	67.97	.45	.29	.81
Q33	73.38	67.53	.58	.46	.81
Q34	73.69	68.84	.41	.21	.81
Q35	73.53	68.45	.57	.42	.81
Q36	73.48	67.57	.60	.44	.81
Q38	73.45	68.28	.55	.48	.81
Q39	73.53	68.11	.63	.51	.81
Q40	73.71	66.62	.57	.43	.81
Q41	74.43	69.82	.31	.16	.82
Q43	73.57	68.45	.42	.30	.81
Q44	73.53	67.37	.56	.43	.81
Q45	73.41	67.55	.61	.46	.81

After removing the items (Items 10, 11, and 16) with low item-total correlations (i.e., below .30), Cronbach's alpha reliability, and Cronbach's alpha based on standardized items became .86 and .87, respectively.

Finally, reliability for the interactional justice component was estimated, and the initial result turned out to be .75. Three items were found to have low item-total correlations (i.e., below .30) as a result of checking item-total statistics (Table 4).

Table 4. *The Item-Total Statistics of the Reliability for the Interactional Justice Component*

	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Squared Multiple Correlation	Cronbach's Alpha if Item Deleted
Q02	49.42	31.50	.34	.23	.73
Q06	49.06	31.10	.43	.32	.73
Q07	48.91	30.95	.51	.28	.72
Q09	49.99	35.47	-.03	.07	.78
Q19	49.50	31.21	.38	.17	.73
Q20	48.86	31.96	.23	.17	.75
Q21	48.98	31.79	.42	.26	.73
Q25	49.09	29.72	.53	.41	.71
Q27	48.82	30.48	.58	.47	.71
Q28	48.88	31.09	.48	.36	.72
Q29	48.78	31.22	.52	.38	.72
Q31	48.95	29.92	.50	.38	.72
Q37	49.63	31.85	.21	.06	.76

After removing the three items (Items 9, 20, and 37), Cronbach's alpha reliability and Cronbach's alpha based on standardized items became .81 and .81, respectively.

5.4. Validity Analysis

As reported above, the reliability analysis resulted in the removal of a total of eight items, leaving the questionnaire with 38 items. To validate the developed instrument, two approaches were taken. First, an EFA was run to capture the whole pattern using Principal Axis Factoring (PAF). Subsequently, the resulting pattern was tested against a CFA using covariance-based software, i.e., IBM AMOS (version 24). Following Worthington and Whittaker (2006) and Van Prooijen, van der Kloot (2001) who recommend running EFA and CFA on the same set of data to provide empirical support regarding the data set, the researchers did CFA and EFA on the same sample.

5.4.1. Exploratory Factor Analysis

Thirty-eight items, left after the reliability analysis, were tested for the underlying patterns using EFA, which is useful for specifying items functioning better or worse (Cheng, 2017). First, sample size adequacy was checked by Bartlett's Test of Sphericity and Kaiser-Meyer-Olkin (KMO) Measure of Sampling Adequacy. The sample had a KMO value of .90, which shows the adequacy of the sample size since KMO values above .60 are considered acceptable. Bartlett's test result was significant; thus, the null hypothesis that the items are independent is rejected. Therefore, they were sufficiently correlated. When Bartlett's test is significant, i.e. $p<.05$, one can say that the correlation matrix is significantly different from an identity one, meaning that there is not zero correlation among all items (Field, 2018). Moreover, EFA requires that the matrix of correlations between items should be neither singular (perfect correlations among all items; =>.90) nor an identity one (zero correlation among all items) (Field, 2018).

Having made sure of the adequacy of the sample size, the EFA was run with the PAF extraction. Table 5 presents the explained total variance.

Table 5. *EFA: Explained Total Variance*

Item	Initial Eigenvalues			Extraction Sums of Squared Loadings		
	Total	*% of Variance*	*Cumulative %*	*Total*	*% of Variance*	*Cumulative %*
1	10.633	27.982	27.982	10.095	26.565	26.565
2	2.379	6.260	34.242	1.807	4.755	31.320
3	1.492	3.927	38.169	1.005	2.644	33.964
4	1.363	3.587	41.756	.806	2.120	36.084

5	1.262	3.322	45.078	.662	1.741	37.825
6	1.199	3.155	48.233	.597	1.571	39.396
7	1.107	2.913	51.147	.517	1.361	40.757
8	1.074	2.827	53.974	.495	1.303	42.061
9	1.015	2.672	56.645	.434	1.142	43.202
10	.975	2.565	59.210			
.	.	.	.			
.	.	.	.			
.	.	.	.			
38	.238	.626	100.00			

Extraction Method: Principal Component Analysis.

The PAF showed extraction of nine factors explaining 43.2% of the total variance. However, the total extracted sum of squares loading was above 1 for only 3 factors.

Figure 1 indicates a scree plot of the results. According to the figure, there was a clear cut after the third factor, and then the line had a steady pattern. This also confirms the results of the sum of squares loading, which suggested keeping the three factors.
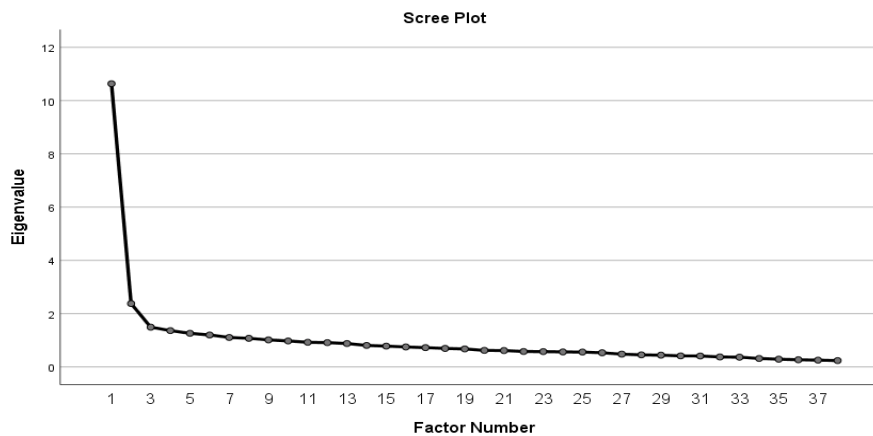

Figure 1. EFA: The Scree Plot

Furthermore, parallel analysis was run on the data using Watkins' (2006) software. The following table shows the results of Watkins' simulation for determining the number of factors to be extracted. The results of Watkins' simulation indicated that three factors should be retained (Table 6).

Table 6. *Parallel Analysis Using Watkins' (2006) Software*

Factors	Criterion	Observed Eigenvalue	Decision
1	**1.740**	**11.070**	**Retain**
2	**1.661**	**2.631**	**Retain**
3	**1.605**	**1.713**	**Retain**
4	1.553	1.532	Drop
5	1.512	1.510	Drop
6	1.470	1.438	Drop
7	1.429	1.322	Drop
8	1.393	1.250	Drop
9	1.359	1.203	Drop
10	1.325	1.173	Drop
11	1.292	1.041	Drop
12	1.263	0.999	Drop
13	1.234	0.981	Drop
14	1.206	0.959	Drop
15	1.176	0.896	Drop
16	1.149	0.860	Drop
17	1.124	0.849	Drop

18	1.097	0.799	Drop
19	1.070	0.796	Drop
20	1.045	0.777	Drop
21	1.022	0.755	Drop
22	0.999	0.710	Drop
23	0.975	0.678	Drop
24	0.953	0.665	Drop
25	0.929	0.649	Drop
26	0.908	0.619	Drop
27	0.884	0.598	Drop
28	0.861	0.577	Drop
29	0.838	0.566	Drop
30	0.816	0.550	Drop
31	0.795	0.515	Drop
32	0.774	0.488	Drop
33	0.751	0.480	Drop
34	0.731	0.457	Drop
35	0.709	0.432	Drop
36	0.687	0.418	Drop
37	0.666	0.382	Drop
38	0.644	0.371	Drop
39	0.624	0.343	Drop
40	0.604	0.337	Drop
41	0.583	0.328	Drop
42	0.561	0.280	Drop
43	0.538	0.273	Drop
44	0.513	0.257	Drop
45	0.485	0.246	Drop
46	0.452	0.226	Drop

Finally, the EFA was re-run with the PAF extraction, Promax rotation, and the three fixed factors. The choice of Promax rotation was made as the inspection of the correlation matrix of the three factors suggested strong correlations among them calling for an oblique rotation, i.e., Promax. Table 7 shows the resulting pattern. According to Boateng et al. (2018), "items with cross-loadings or that appear not to load uniquely on individual factors can be deleted" (p. 11). Thus, based on the results, three items (Items 4, 14, 23) were loaded onto two factors with close estimates. None of the loadings were above .4, though. Two other items (Q41 and Q42) had positive loading to one factor and negative loading to another. In extracting the factors, as recommended by Boateng et al. (2018), Raykov and Marcoulides (2011), and Nunnally (1978), we only kept those with loadings above .4. In the same vein, Field (2018) quoted Stevens (2002) who believed that factor loadings greater than .4 should be interpreted. This solved the issue of double saturation as all of the items with those characteristics had either one or no loadings above .4. Thus, as a result of the PAF, 20 questions were kept.

Table 7. *EFA: The Varimax Rotated Pattern Matrix*

Item	Factor		
	1	2	3
Q02	.607	.011	-.251
Q03	.230	.223	.092
Q04	*.315*	*.323*	-.092
Q05	.003	.441	.098
Q06	.684	-.170	-.037
Q07	.467	.134	.012
Q12	.343	.222	-.228
Q13	.342	-.073	.185
Q14	*.309*	.139	*.328*
Q15	-.040	-.109	.759
Q17	.117	.075	.358
Q18	.198	.117	.205
Q19	.320	.243	-.186

Q21	.491	.074	-.029
Q22	.314	-.068	.013
Q23	*342*	.168	*331*
Q24	.346	.208	-.038
Q25	.595	-.077	.065
Q26	-.312	.114	.602
Q27	.528	.135	.028
Q28	.314	.195	.130
Q29	.434	.184	.098
Q30	.029	-.237	.792
Q31	.446	.109	.004
Q32	.093	.351	.132
Q33	.094	.478	.166
Q34	.334	.147	.038
Q35	.272	.380	.069
Q36	.141	.510	.088
Q38	.273	.222	.264
Q39	.201	.457	.094
Q40	.041	.725	-.197
Q41	-.081	.649	-*381*
Q42	-*326*	.223	.462
Q43	.107	.342	.097
Q44	.183	.532	-.082
Q45	-.144	.780	.071
Q46	.280	.248	.053

Extraction Method: Principal Axis Factoring.

Rotation Method: Promax with Kaiser Normalization.

a. Rotation converged in 5 iterations.

Based on triangulating the content of the left items and their loaded factors, Factor 1, Factor 2, and Factor 3 were named *Interactional Justice*, *Procedural Justice*, and *Distributive Justice*, respectively.

5.4.2. Confirmatory Factor Analysis

At this stage, three sets of CFA were run based on the pattern obtained from EFA. Three CFAs (one with a one-factor solution, one with a three-factor, and one with a three-factor modified based on modification indices) were run with the maximum likelihood method in IBM AMOS (version 24). Here, we will first report the results for the three-factor solution models (before and after modification) and then, a comparison of the model fits will be done between these models and the one-factor solution model to reach the best choice.

Before running the CFA, the sample size was checked for its adequacy through various rules-of-thumb; Boomsma (1982) recommended the use of at least 150 observations, while Bentler and Chou (1987) set the rule of five to 10 observations per estimated parameter. Kline (2016) recommended 20 observations per estimated parameter. Having a sample of 373 participants was large enough to run a CFA on the three parameters and the 20 items. Moreover, to decide the appropriate choice for running CFA (choosing between parametric and non-parametric maximum likelihood), the normality of the 20 items in the model was checked (Table 8). The results of skewness (ranging from -2.038 to -.297) and kurtosis (ranging from -.271 to 4.578) indicated normality for all items. As stated by Brown (2006), when employing SEM, kurtosis is appropriate from a range of -10 to $+10$, and skewness is acceptable within a range of -3 and $+3$. Therefore, the parametric maximum likelihood was used in running CFA.

Table 8. *Checking the Normality of Distributions*

Item	N	Mean	SD	Skewness	Kurtosis
Q02	373	3.82	.947	-.545	-.136
Q05	373	4.33	.837	-1.264	1.502
Q06	373	4.18	.877	-.857	.182

Q07	373	4.33	.797	-1.201	1.657
Q15	373	4.35	1.078	-1.706	1.941
Q21	373	4.26	.771	-.903	.698
Q25	373	4.15	.947	-1.011	.547
Q26	373	4.04	1.251	-1.186	.261
Q27	373	4.42	.778	-1.269	1.220
Q29	373	4.47	.742	-1.393	1.821
Q30	373	4.35	1.210	-1.678	1.366
Q31	373	4.29	.951	-1.474	1.923
Q33	373	4.53	.795	-2.038	4.587
Q36	373	4.43	.775	-1.566	3.004
Q39	373	4.37	.690	-1.088	1.834
Q40	373	4.19	.891	-.979	.541
Q41	373	3.47	.966	-.297	-.014
Q42	373	3.85	1.167	-.811	-.271
Q44	373	4.37	.838	-1.673	3.532
Q45	373	4.49	.764	-1.831	4.003

Table 9 shows the standardized and unstandardized estimates results of the three-factor solution model. The items with non-significant unstandardized estimates and standardized estimates below .4 had to be omitted (Items 2 and 41).

Table 9. *The Unstandardized and Standardize Estimates of the Initial CFA Model*

			Unstandardized				Standardized
			Estimate	*S.E.*	*C.R.*	*P*	*Estimate*
Q02	<---	Interactional	1.000				**.340**
Q06	<---	Interactional	1.254	.229	5.465	.000	.528
Q07	<---	Interactional	1.254	.222	5.638	.000	.581
Q21	<---	Interactional	1.067	.198	5.404	.000	.511
Q25	<---	Interactional	1.433	.257	5.570	.000	.559
Q27	<---	Interactional	1.316	.229	5.756	.000	.625
Q29	<---	Interactional	1.338	.229	5.848	.000	.666
Q31	<---	Interactional	1.538	.271	5.683	.000	.597
Q05	<---	Procedural	1.000				.470
Q33	<---	Procedural	1.352	.166	8.123	.000	.669
Q36	<---	Procedural	1.304	.161	8.085	.000	.662
Q39	<---	Procedural	1.169	.144	8.110	.000	.666
Q40	<---	Procedural	1.500	.186	8.086	.000	.662
Q41	<---	Procedural	.742	.153	4.864	.000	**.302**
Q44	<---	Procedural	1.355	.171	7.935	.000	.636
Q45	<---	Procedural	1.372	.165	8.319	.000	.706
Q15	<---	Distributive	1.000				.677
Q26	<---	Distributive	.973	.116	8.385	.000	.568
Q30	<---	Distributive	1.229	.132	9.274	.000	.742
Q42	<---	Distributive	.683	.103	6.654	.000	.427

After removing the two items, the modifications suggested by the software with the threshold of 10 were considered, and those with no conflict to the literature and positive par change in the model were applied. Figure 2 depicts the modified CFA model with standardized estimates.
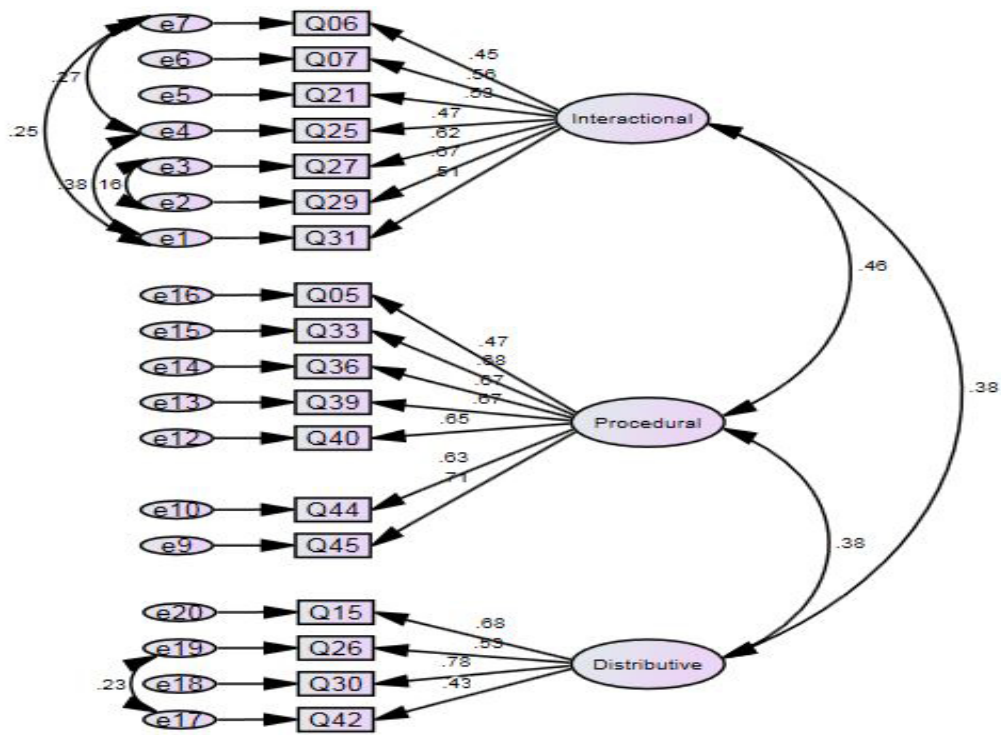
Figure 2. The Modified CFA Model with Standardized Estimates

Having applied the modifications, the researchers checked the goodness of fit of the model. As recommended by Hu and Bentler (1999), some criteria need to be met for ensuring the goodness of fit of the model. These criteria, as well as the values obtained from the three models of CFA, are reported below in Table 10.

Table 10. *Evaluation of the CFA Goodness of Fit*

Criteria	Excellent Threshold	Models			Best Model Fit
		One-Factor	*Three-Factor*	*Modified*	
CMIN		627.995	364.324	267.678	
Df		135	132	127	
CMIN/df	Between 1 and 3	4.652	2.760	2.108	Modified
RMSEA	<.06	.099	.069	.055	Modified
SRMR	<.08	.087	.059	.049	Modified
CFI	>.95	.748	.881	.928	Modified
PClose	> .05	Not estimated	Not estimated	.198	

The excellent threshold values in Table 10 are reported according to Hu and Bentler (1999). They also argue that values exceeding the range may be considered acceptable if they fall within a certain range (between 3 and 5 for CMIN/df; between .06 and .08 for RMSEA; between .08 and 1 for SRMR, and between .9 and .95 for CFI). According to Table 10, the one-factor solution resulted in terrible RMSEA and CFI indices while CMIN/df and SRMR can be considered acceptable. The three-factor solution resulted in an acceptable CMIN/df and SRMR, acceptable RMSEA, and terrible CFI. However, the three-factor modified solution showed excellent indices of model fit in all cases except for CFI, which had an acceptable value. Therefore, the modified model was considered the best choice.

Next, as presented in Table 11, the researchers checked the composite reliability and discriminant validity of each factor. All the factors had values above .70, showing their acceptable reliability. Additionally, the square root of the average variance extracted (AVE) (the bold values in the table) was above the inter-correlations of the factors, showing discriminant validity based on Fornell and Larcker's recommendation (1981).

Table 11. *The Composite Reliability and Discriminant Validity of the Factors*

	CR	Fornell-Larcker Criterion		
		Distributive	*Procedural*	*Interactional*
Interactional	**.749**	**.550**		
Procedural	**.829**	.462	**.642**	
Distributive	**.705**	.382	.385	**.607**

Table 12 presents the finalized TCJS, which include 18 items, measuring the three main components of interactional, procedural, and distributive classroom justice.

Table 12. *The Finalized 18-Item TCJS*

1. If a student criticizes me, I will think about my behavior or practice instead of dismissing the student.
2. I make timely communication of my expectations with students at the beginning of the semester.
3. I provide my students with sufficient and honest information regarding the criteria that I employ in grading them.
4. I grade students based on their achievements.
5. The exams in my class include enough explanations and instructions for students to show what they know and what they can do.
6. I specify my expectations, syllabus, and grading criteria at the beginning of the semester.
7. I provide equal attention and help to both high- and low-achievers in my class.
8. I have a caring and supportive relationship with my students.
9. I am sensitive to my students' feelings, opinions, and rights.
10. I provide students in the class with equal access to information on exam materials.
11. I adequately inform students of my class attendance policy.
12. If I teach a topic incorrectly in class, I will accept and attempt to provide correct information.
13. I always set equal opportunity and time to students to participate in classroom discussions.
14. I set reasonable expectations for my students' performance.
15. I give my students the opportunity to express their views and feelings regarding the scheduling of homework and topics I plan to set in the course.
16. I do not have a favorite student who I treat differently from other students in class.
17. I explain everything clearly, so students understand.
18. I allow my students to express their concerns regarding the attendance policy.

The Three Components Underlying the Scale: (1) Distributive Justice (Items 4, 7, 10, 16); Procedural Justice (Items 1, 12, 13, 14, 15, 17, 18); Interactional Justice (Items 2, 3, 5, 6, 8, 9, 11). Responses range from 1 "never", 2 "rarely", 3 "often", 4 "usually", to 5 "always".

6. Discussion

This research undertaking aimed to develop and validate an instrument for measuring Iranian EFL instructors' justice behaviors. To this aim, a draft of the questionnaire including 46 items was developed after thoroughly reviewing the literature, interviewing experts in the field, and scrutinizing the existing scales. The scale items reflected the comprehensive framework outlined by Estaji and Zhaleh (2021a) including the interactional, procedural, and distributive classroom justice dimensions. Accordingly, distributive classroom justice was realized through the principles of need, equality, and equity, procedural classroom justice was explained in terms of bias suppression, ethicality, accuracy, consistency, transparency, voice, correctability, and reasonableness principles, and interactional classroom justice was described through the principles of caring, propriety, respect, timeliness, truthfulness, and adequacy/justification. Furthermore, the questionnaire items reflected the explanation of justice dimensions and principles in classroom domains of assessment, interactions, learning, and teaching.

Psychometric properties of this scale were examined in four steps: (1) doing a pilot testing to identify the unexpected problems in the administration of the questionnaire, reaching a priori estimation of reliability, and identifying items that could lower reliability in the main administration; (2) analyzing reliability; (3) conducting EFA to determine the number of factors that could best explain the scale; and (4) performing CFA to test the resulted patterns. The outcomes of the goodness-of-fit indices on the modified model demonstrated an acceptable model-to-data fit and approved factor structure of the scale. In this study, the initial item pool of the scale including 46 items was reduced, and the finalized included 18 items. In this respect, during the reliability analysis stage where item-total statistics was inspected for each

component, eight items were omitted as their item-total correlations were below .30 (Field, 2018; Pallant, 2016). Subsequently, during the EFA analysis stage through running PAF extraction, parallel analysis (Watkins, 2006), as well as Promax rotation, it was found that three factors should be retained for the scale. Accordingly, 18 items were omitted at this stage since they had cross-loadings or loadings below .40 (Boateng et al., 2018; Field, 2018; Nunnally, 1978; Raykov & Marcoulides, 2011; Stevens, 2002). Finally, during CFA, two other items with non-significant unstandardized estimates and standardized estimates below .4 had to be omitted. Notwithstanding this reduction in the number of scale items from 46 to 18, acceptable composite reliability and discriminant validity values were obtained for the three factors of the finalized TCJS. Thus, the finalized TCJS can be regarded as an efficient instrument for measuring perceptions that EFL instructors in Iran hold toward their classroom justice behaviors.

The finalized scale included 18 items in terms of the factors of distributive, procedural, and interactional justice. As pointed out by Hwang and Lee (2017), researchers are required to compare the scale validation results of their study with those of the previous pertinent studies to ensure the soundness of their obtained factors in light of the frameworks and theories in the literature. Accordingly, in the present study, it was found that the findings are in agreement with those of the previous studies reporting distributive, interactional, and procedural justice as the main dimensions of classroom justice (Berti et al., 2010; Chory et al., 2017; Di Battista et al., 2014; Rasooli et al., 2019). The results also confirmed that the Western social psychology theories of justice in education (e.g., Donat et al., 2018; Kazemi, 2008; Sabbagh & Resh, 2016) can be extended to the Iranian EFL context. This shows that teachers share some common perceptions about classroom justice since certain justice principles probably remain the same, notwithstanding the varieties in the educational domains, contexts, or cultures.

Aside from the mentioned commonalities of the current research findings with the previous studies, which showed that the three TCJS components were compatible with the previous literature in terms of the three overarching dimensions of classroom justice (i.e., procedural, interactional, and distributive), it should be, however, pinpointed that some individual items of TCJS were unique to the Iranian EFL instructional context. More particularly, during the TCJS item pool development stage, items 2, 7, 8, 11, 14, 15, and 18 were developed based on the results of the interview with the Iranian EFL teachers, thus, revealing context-specific concerns about justice raised by Iranian EFL teachers. It is justifiable to explain this finding by noting the uniqueness of L2 learning and teaching processes which are different from those of other subject matters since the attainment of language as a communication-based and relational process requires good rapport, interpersonal relationships, and mutual understanding between the language teacher and students (Mercer & Dörnyei, 2020). Therefore, it is more likely that L2 teachers implement more justice principles to create a positive and effective language learning environment for their students than other teachers (Estaji & Zhaleh, 2021a).

The TCJS developed in this study is a more comprehensive tool for assessing the perceptions of classroom justice in comparison with the existing scales as it represents justice in different subdomains of the classroom such as grade distribution, exam content, teacher care, syllabus design, and implementation, teacher help, class attendance policy, classroom participation opportunity, and teacher explanations. This finding also empirically supports this argument that teachers are responsible for enacting fairness in every single aspect of their instructional practice, and being fair in one classroom domain (e.g., assessment) to the disregard of other equally important domains (e.g., teaching, interactions, and learning) would not provide a comprehensive picture of teacher classroom justice behavior (Horan et al., 2010). As articulated by Horan and Myers (2009), instructors' overall fairness mindset is shaped through the unique contribution and interplay of the assessment, teaching, interactions, and learning classroom elements. Previous studies have approved this notion by showing that students and teachers make fairness conceptualizations within all these classroom elements (e.g., Estaji & Zhaleh, 2021a, 2021b, 2022; Čiuladienė & Račelytė, 2016; Rasoolie et al., 2019; Robbins & Jeffords, 2009).

7. Conclusion and Implications

Based on the present study results, aiming to explore the underlying components of teacher classroom justice in the Iranian EFL context, it can be concluded that the Western social psychology theories of classroom justice in general and the main dimensions of distributive, interactional, and procedural justice, can be extended to the Iranian L2 education context. More interestingly, the unique nature of L2 learning and teaching demands more teacher-student mutual trust and understanding, respect, honesty, positive interpersonal relationships, and caring (Mercer & Dörnyei, 2020) which highlights the teachers' serious attention to every principle of justice in the instructional context (Estaji & Zhaleh, 2021a).

Some of the items developed from the interview results were retained in the finalized scale after CFA and EFA. These items reflect the context-specific nature of the TCJS, which was developed and validated in the current study as they attend to the distinctive features of the Iranian EFL education context. Overall, the TCJS developed based on this conceptualization of classroom justice is a good and valid measure of classroom justice perceptions of EFL instructors in Iran.

As pointed out before, questionnaires have been the most prevailing instrument for measuring the students' or teachers' perceptions of classroom justice in general education (Rasooli, 2020). However, no comprehensive justice scale existed in the L2 language education domain. Therefore, as the first empirical stride toward developing and validating a TCJS questionnaire in L2 education, the present study has useful theoretical and pedagogical implications for this domain. Theoretically, the obtained results add invaluable insights to the expanding line of research on classroom justice and the social psychology theory of justice as a whole. Thus, it is hoped that after the development of the TCJS in the present study, more survey-based studies be conducted in this area in the future to expand the literature on classroom justice in L2 education and expand this under-represented line of research. By employing the TCJS, future studies can examine the associative or causal link of teacher-classroom justice with an array of teacher-related factors, such as work engagement, job satisfaction, reflection, burnout, well-being, or effectiveness, or student-related factors like motivation, anxiety, learning, enjoyment, autonomy, or achievement in the EFL context.

Practically, the outcomes of this study can enlighten the practice and mindset of language education researchers, teacher educators, teacher professional development course designers, and EFL instructors. Regarding EFL instructors, the findings can expand their knowledge and awareness of classroom justice and its underlying components. Teacher educators can benefit from the results by teaching the concept of classroom justice to pre- and in-service teachers in teacher training courses or teacher education programs and equip them with useful recommendations and strategies about how to apply justice dimensions and principles in their practices in classroom domains of assessment, interactions, teaching, and learning. In the same vein, teacher development course designers can incorporate the components of TCJS in their programs. Finally, language education researchers can employ the developed and validated questionnaire of this study to explore EFL instructors' perceived level of classroom justice.

The current research has some limitations. First, some actions such as choosing participants from different genders, majors, teaching experience, and educational levels were taken to augment the sample-to-population extrapolation of the data. Nevertheless, the questionnaire was validated on a group of around 400 Iranian EFL teachers, and thus, the findings should be cautiously generalized to the statistical population of EFL instructors worldwide. Second, the TCJS only measures EFL teachers' classroom justice perceptions. As teachers and learners both play key roles in the experience of justice in the classroom, it is necessary that the learners' perceptions of justice be examined in concomitance with the teachers' perceptions. To this aim, future studies can revalidate the TCJS on EFL students. Finally, the present scale was developed in an EFL context. In the future, researchers can examine other educational contexts with second and foreign languages other than English to cater for different language education contexts.

References

Adams, J. S. (1965). Inequity in social exchange. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (pp. 267-299). Academic Press.

Bentler, P. M., & Chou, C. P. (1987). Practical issues in structural modeling. *Sociological Methods & Research, 16*(1), 78–117.

BERA. (2011). *Ethical guidelines for educational research*. Retrieved April 4, 2016, from http://content.yudu.com/Library/A2xnp5/Bera/resources/index.htm?referrerUrl=http://free.yudu.com/item/details/2023387/Bera

Berti, C., Molinari, L., & Speltini, G. (2010). Classroom justice and psychological engagement: Students' and teachers' representations. *Social Psychology of Education, 13*(4), 541–556. https://doi.org/10.1007/s11218-010-9128-9

Bies, R. J., & Moag, J. F. (1986). Interactional justice: Communication criteria of fairness. In R. J. Lewicki, B. H. Sheppard, & M. H. Bazerman (Eds.), *Research on negotiation in organizations* (pp. 43–55). JAI Press.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: A Primer. *Frontiers in Public Health, 6*, 149. https://doi.org/10.3389/fpubh.2018.00149

Boomsma, A. (1982). Robustness of LISREL against small sample sizes in factor analysis models. In K. G. Joreskog & H. Wold (Eds.), *Systems under indirect observation: Causality, structure, prediction* (Part 1, pp. 149–173). North Holland.

Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. The Guilford Press.

Cheng, Y. (2017). Development and preliminary validation of four brief measures of L2 language-skill-specific anxiety. *System, 68*, 15–25. https://doi.org/10.1016/j.system.2017.06.009

Chory, R. M. (2007). Enhancing student perceptions of fairness: The relationship between instructor credibility and classroom justice. *Communication Education, 56*(1), 89–105. https://doi.org/10.1080/03634520600994300

Chory, R. M., Horan, S. M., & Houser, M. L. (2017). Justice in the higher education classroom: Students' perceptions of unfairness and responses to instructors. *Innovative Higher Education, 42*(4), 321–336. https://doi.org/10.1007/s10755-017-9388-9

Chory, R. M., Zhaleh, K., & Estaji, M. (2022). Perceptions of instructor injustice in COVID-19-imposed online courses: EFL students' perceptions and experiences in focus. *Communication Quarterly, 70*(5), 469-494. https://doi.org/10.1080/01463373.2022.2090265

Chory-Assad, R. M., & Paulsel, M. L. (2004). Classroom justice: Student aggression and resistance as reactions to perceived unfairness. *Communication Education, 53*(3), 253–273. https://doi.org/10.1080/0363452042000265189

Čiuladienė, G., & Račelytė, D. (2016). Perceived unfairness in teacher-student conflict situations: Students' point of view. *Polish Journal of Applied Psychology, 14*(1), 49–66. https://doi.org/10.1515/pjap-2015-0049

Cropanzano, R., & Greenberg, J. (1997). Progress in organizational justice: Tunneling through the maze. In C. L. Cooper & I. T. Robertson (Eds.), *International Review of Industrial and Organizational Psychology* (pp. 317–372). John Wiley & Sons.

Cropanzano, R., Fortin, M., & Kirk, J. F. (2015). How do we know when we are treated fairly? Justice rules and fairness judgments. In M. R. Buckley, A. R. Wheeler, & J. R. B. Halbesleben (Eds.), *Research in personnel and human resources management* (pp. 279–350). Emerald Group Publishing Limited.

Dalbert, C. (2013). Die Bedeutung schulischen Gerechtigkeitserlebens fu¨r das subjektive Wohlbefinden in der Schule [The meaning of justice experiences for subjective well-being at school]. In C. Dalbert (Ed.), *Gerechtigkeit in der Schule* (pp. 127–143). Springer VS.

Dalbert, C., & Stoeber, J. (2006). The personal belief in a just world and domain-specific beliefs about justice at school and in the family: A longitudinal study with adolescents. *International Journal of Behavioral Development, 30*(3), 200–207. https://doi.org/10.1177/0165025406063638

Derakhshan, A., Coombe, C., Zhaleh, K., & Tabatabaeian, M. (2020). Examining the roles of continuing professional development needs and views of research in English language teachers' success. *The Electronic Journal for English as a Second Language (TESL-EJ), 24*(3), 1–27. http://www.tesl-ej.org/pdf/ej95/a2.pdf

Derakhshan, A., Doliński, D., Zhaleh, K., Janebi Enayat, M., & Fathi, J. (2022). A mixed-methods cross-cultural study of teacher care and teacher-student rapport in Iranian and Polish University students' engagement in pursuing academic goals in an L2 context. *System, 106*, 102790. https://doi.org/10.1016/j.system.2022.102790

Deutsch, M. (1975). Equity, equality, and need: What determines which value will be used as the basis of distributive justice? *Journal of Social Issues, 31*(3), 137–149. https://doi.org/10.1111/j.1540-4560.1975.tb01000.x

Di Battista, S., Pivetti, M., & Berti, C. (2014). Engagement in the university context: Exploring the role of a sense of justice and social identification. *Social Psychology of Education: An International Journal, 17*(3), 471–490. https://doi.org/10.1007/s11218-014-9255-9

Shahid Chamran University of Ahvaz

Donat, M., Gallschütz, C., & Dalbert, C. (2018). The relation between students' justice experiences and their school refusal behavior. *Social Psychology of Education, 21*(2), 447–475. https://doi.org/10.1007/s11218-017-9423-9

Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing.* Lawrence Erlbaum.

Ehrhardt, N., Pretsch, J., Herrmann, I., & Schmitt, M. (2018). Observing justice in the primary school classroom. *Zeitschrift Für Erziehungswissenschaft, 19*(1), 157–190. https://doi.org/10.1007/s11618-015-0664-0

Estaji, M., & Zhaleh, K. (2021a). Exploring Iranian teachers' perceptions of classroom justice and its dimensions in EFL instructional contexts. *Language Related Research, 12*(3), 277–314. https://doi.org/10.29252/LRR.12.3.10

Estaji, M., & Zhaleh, K. (2021b). Teachers' perceptions, experiences, and challenges of incorporating justice in English as a foreign language classrooms. *Polish Psychological Bulletin, 52*(3), 251–269. https://doi.org/10.24425/ppb.2021.137889

Estaji, M., & Zhaleh, K. (2022a). Experiences, causes, and solutions for injustice in Covid-19 online classes: Evidence from university EFL students in Iran. *LEARN Journal: Language Education and Acquisition Research Network, 15*(2), 436-467. https://so04.tci-thaijo.org/index.php/LEARN/article/view/259936

Estaji, M., & Zhaleh, K. (2022b). Iranian EFL teachers' expectations and perceptions about teacher classroom justice training. *TEFLIN Journal, 33*(2), 272-291. http://doi.org/10.15639/teflinjournal.v33i2/272-291

Estaji, M., & Zhaleh, K. (2022c). The enactment of classroom justice through explicit instruction: Deciphering the changes in English as a foreign language teachers' perceptions and practices. *Frontiers in Psychology, 13*, 821763. https://doi.org/10.3389/fpsyg.2022.821763

Field, A. (2018). *Discovering statistics using IBM SPSS statistics* (5th ed.). SAGE Publications.

Fornell, C. & Larcker, D. F. (1981). Evaluating structural equation models with unobservable variables and measurement error. *Journal of Marketing Research*, 18(1), 39–50. https://doi.org/10.1177/002224378101800104

Gao, L. X., & Zhang, L. J. (2020). Teacher learning in difficult times: Examining foreign language teachers' cognitions about online teaching to tide over COVID-19. *Frontiers in Psychology, 11*, 2396. https://doi.org/10.3389/fpsyg.2020.549653

Gasser, L., Grütter, J., Buholzer, A., & Wettstein, A. (2018). Emotionally supportive classroom interactions and students' perceptions of their teachers as caring and just. *Learning and Instruction, 54*, 82–92. https://doi.org/10.1016/j.learninstruc.2017.08.003

Ghaedsharafi, A., Yamini, M., & Dehghan, F. (2019). Investigating predictors of high school students' negative attitudes towards learning English by developing, validating, and running a questionnaire. *Journal of Research in Applied Linguistics, 10*(2), 136–158. https://doi.org/10.22055/RALS.2019.14721

Gorard, S. (2012). Experiencing fairness at school: An international study. *International Journal of Educational Research, 53*(3), 127–137. https://doi.org/10.1016/j.ijer.2012.03.003

Gouveia-Pereira, M., Vala, J., Palmonari, A., & Rubini, M. (2003). School experience, relational justice, and legitimation of institutional authorities. *European Journal of Psychology of Education, 18*(3), 309–325. https://doi.org/10.1007/BF03173251

Horan, S. M., & Myers, S. (2009). An exploration of college instructors' use of classroom justice, power, and behavior alteration techniques. *Communication Education, 58*, 483–496. http://doi.org/10.1080/03634520903055981

Horan, S. M., Chory, R. M., & Goodboy, A. K. (2010). Understanding students' classroom justice experiences and responses. *Communication Education, 59*(4), 453–474. https://doi.org/10.1080/03634523.2010.487282

Hu, L., & Bentler, P. M. (1999). Cut-off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *SEM, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Shahid Chamran University of Ahvaz

Hwang, M., & Lee, H.-K. (2017). Development and validation of the English writing strategy inventory. *System, 68*, 60–71. https://doi.org/10.1016/j.system.2017.06.014

Kazemi, A. (2008). Social psychology of justice: Origins, central issues, recent developments, and future directions. *Nordic Psychology, 60*(3) 209–234. https://doi.org/10.1027/1901-2276.60.3.209

Kazemi, A. (2016). Examining the interplay of justice perceptions, motivation, and school achievement among secondary school students. *Social Justice Research, 29*(1), 103–118. https://doi.org/10.1007/s11211-016-0261-2

Kline, R. B. (2016). *Principles and practice of structural equation modeling* (4th ed.). The Guilford Press.

Lizzio, A., Wilson, K., & Hadaway, V. (2007). University students' perceptions of a fair learning environment: A social justice perspective. *Assessment & Evaluation in Higher Education, 32*(2), 195–213. https://doi.org/10.1080/02602930600801969

Mameli, C., Biolcati, R., Passini, S., & Mancini, G. (2018). School context and subjective distress: The influence of teacher justice and school-specific well-being on adolescents' psychological health. *School Psychology International, 39*(5), 1–17. https://doi.org/10.1177/0143034318794226

Mameli, C., Caricati, L., & Molinari, L. (2020). That's not fair! The effects of teacher justice and academic achievement on Self and Other's resistant agency. *British Journal of Educational Psychology*, *90*(4), 933–947. https://doi.org/10.1111/bjep.12335

Mercer, S., & Dörnyei, Z. (2020). *Engaging language learners in contemporary classrooms*. Cambridge University Press.

Mirshojaee, S. B., Sahragard, R., Razmjoo, S. A., & Ahmadi, A. (2019). Iranian language teachers' passion for the profession: A qualitative study. *Journal of Research in Applied Linguistics, 10*(2), 44–69. https://doi.org/10.22055/RALS.2019.14717

Nassaji, H. (2020). Good qualitative research. *Language Teaching Research, 24*(4), 427–431. https://doi.org/10.1177/1362168820941288

Nunnally J. C. (1978). *Psychometric theory*. McGraw-Hill.

Pallant, J. (2016). *SPSS survival manual: A step by step guide to data analysis using SPSS program* (6th ed.). McGraw-Hill Education.

Pnevmatikos, D., & Trikkaliotis, I. (2012). Procedural justice in a classroom where teacher implements differentiated instruction. In D. Alt & R. Reingold (Eds.), *Changes in teachers' moral role. From passive observer to moral and democratic leadership* (pp. 155–163). Sense Publishers.

Rasooli, A. (2020, 29 May-4 June). *Rethinking survey validation processes: Lessons learned from a systematic review of classroom justice instruments* [Paper Presentation]. Annual Conference of the Canadian Society for the Study of Education, Western University, Canada.

Rasooli, A., DeLuca, C., Rasegh, A., & Fathi, S. (2019). Students' critical incidents of fairness in classroom assessment: An empirical study. *Social Psychology of Education, 22*(3), 701–722. https://doi.org/10.1080/0969594X.2019.1593105

Rasooli, A., Zandi, H., & DeLuca, C. (2018). Re-conceptualizing classroom assessment fairness: A systematic meta-ethnography of assessment literature and beyond. *Studies in Educational Evaluation, 56,* 164–181. https://doi.org/10.1016/j.stueduc.2017.12.008

Raykov, T., & Marcoulides, G. A. (2011). *Introduction to psychometric theory*. Routledge, Taylor & Francis Group.

Resh, N., & Sabbagh, C. (2016). Justice and education. In C. Sabbagh & M. Schmitt (Eds.), *Handbook of social justice theory and research* (pp. 349–367). Springer.

Resh, N., & Sabbagh, C. (2017). Sense of justice in school and civic behavior. *Social Psychology of Education, 20*(2), 387–409. https://doi.org/10.1007/s11218-017-9375-0

Robbins, T. L., & Jeffords, B. C. (2009). Practicing what we preach: Justice and ethical instruction in management education. *Ethics and Education, 4*(1), 93-102. https://doi.org/10.1080/17449640902861562

Shahid Chamran University of Ahvaz

Rodrigues, I. B., Adachi, J, D., Beattie, K. A., & MacDermid, J. C. (2017). Development and validation of a new tool to measure the facilitators, barriers and preferences to exercise in people with osteoporosis. *BMC Musculoskeletal Disorders, 18*(1), 540. https://doi.org/10.1186/s12891-017-1914-5

Roffey, S. (2011). *Changing behaviour in schools: Promoting positive relationships and well-being*. SAGE.

Sabbagh, C., & Resh, N. (2014). Citizenship orientations in a divided society: A comparison of three groups of Israeli junior-high students—secular Jews, religious Jews, and Israeli Arabs. *Education, Citizenship, and Social Justice, 9*(1), 34–54. https://doi.org/10.1177/1746197913497662

Sabbagh, C., & Resh, N. (2016). Unfolding justice research in the realm of education. *Social Justice Research, 29*(1), 1–13. https://doi.org/10.1007/s11211-016-0262-1

Sonnleitner, P., & Kovacs, C. (2020). Differences between students' and teachers' fairness perceptions: Exploring the potential of a self-administered questionnaire to improve teachers' assessment practices. *Frontiers in Education, 5*, 1–14. https://doi.org/10.3389/feduc.2020.00017

Stevens, J. P. (2002). *Applied multivariate statistics for the social sciences* (4th ed.). Erlbaum.

Tabachnick, B. G., & Fidell, L. S. (2018). *Using multivariate statistics* (7th ed.). Pearson.

Thibaut, J. W., & Walker, L. (1975). *Procedural justice: A psychological analysis*. Lawrence Erlbaum Associates.

Tripp, T. M., Jiang, L., Olson, K., & Graso, M. (2019). The fair process effect in the classroom: Reducing the influence of grades on student evaluations of teachers. *Journal of Marketing Education, 41*(3), 1–12. https://doi.org/10.1177/0273475318772618

Tyler, T. R. (1987). Procedural justice research. *Social Justice Research, 1*(1), 41–65. https://doi.org/10.1007/BF01049383

Van Prooijen, J.-W., & van der Kloot, W. A. (2001). Confirmatory analysis of exploratively obtained factor structures. *Educational and Psychological Measurement, 61*(5), 777–792. https://doi.org/10.1177/00131640121971518

Watkins, M. W. (2006). Determining parallel analysis criteria. *Journal of Modern Applied Statistical Methods, 5*(2), 344–346.

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist 34*(6), 806–838. https://doi.org/10.1177/0011000006288127

Zhaleh, K., Estaji, M., & Berti, C. (2022). Conceptualizing classroom justice in second/foreign language education: past, present, & future directions of research & practice. *Journal of Critical Applied Linguistics Studies (JCALS), 1*(1), 32-49. https://doi.org/10.22034/jcals.1.1.32

Zhaleh, K., Ghonsooly, B., & Pishghadam, R. (2018). Effects of conceptions of intelligence and ambiguity tolerance on teacher burnout: A case of Iranian EFL teachers. *Journal of Research in Applied Linguistics, 9*(2), 118-140. https://doi.org/10.22055/RALS.2018.13796

Appendices

Appendix A. *Demographic Information*

Gender:

- Male
- Female

Age:

- Less than 20
- 20-29
- 30-39

- 40-49

- 50 or more

Last academic degree obtained

- Diploma

- Associates of Art

- Bachelor of Art

- Master of Art

- PhD

- Other

Province where you are currently teaching: …………………..

Major

- English Language and Literature

- Teaching English as a Foreign Language (TEFL)

- Applied Linguistics

- English Translation

- Linguistics

- TESOL

- Majors other than English

Years of experience as an English language teacher

- 0-4

- 5-9

- 10-14

- 15-19

- 20-24

- 25 or more

The proficiency level(s) you are currently teaching

- Beginner

- Pre-intermediate. Intermediate. Advanced

- Proficient

The age range of students you are currently teaching.

- Children

- Teenagers

- Adults

Appendix B. *The Interview Prompts Employed in Interview With Experts in the Field*

1. Is the idea of being a just teacher important to you in your daily work? If so, why? Lf not, why not?

2. How do you define the teacher classroom justice behavior generally in your own tenns?

3. In what classroom aspects do you think teacher justice can be implemented?

4. Explain how teachers can employ justice when distributing educational outcomes such agrades, feedback, reward, help, time, or punislunent among their students.
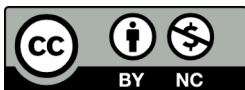
5. Explain how teachers can incorporate justice when enacting classroom procedures and policies.

6. Explain how teachers can apply justice in their interpersonal relationships with students and communication of information to them.

7. As an EFL teacher, what types of teacher classroom behaviors do you consider to be just?

8. Can you remember and explain some specific situations that you behaved in a just way toward your students?

9. As an EFL teacher, what types of teacher classroom behaviors do you consider to be unjust?

10. Can you remember and explain some specific situations that you behaved in an unjust way toward your students?

11. Do you evaluate yourself as a just or unjust EFL teacher? Why?

12. What are some of the predicaments or obstacles that an EFL teacher may face when trying to be just in the classroom?

13. What challenges or obstacles have you yourself ever experienced when trying to be fair in your classroom?

Appendix C. *The 46 Items of the Questionnaire, and the Dimension, Principle, and Domain that Each Item Belongs To*

Instruction: The following items assess the degree to which you, as an English language teacher, perceive your performance to be fair and just in the instructional context. For each item, please answer using the following scale: Never (1), Rarely (2), Often (3), Usually (4), to Always (5).

1) I give reward to students considering their study commitment, effort, and contribution	Distributive	Equity	Learning
2) I pay attention to my students' learning abilities, styles, and exceptionalities.	Interactional	Caring	Learning
3) I maintain appropriate eye contact with each student in my class.	Distributive	Equality	Interaction
4) When I make a decision about my students, I seek to understand everything that has happened in a particular situation.	Procedural	Accuracy	Teaching
5) If a student criticizes me, I will think about my behavior or practice instead of dismissing the student.	Procedural	Correctability	Interaction
6) I make timely communication of my expectations with students at the beginning of the semester.	Interactional	Timeliness	Teaching
7) I provide my students with sufficient and honest information regarding the criteria that I employ in grading them.	Interactional	Truthfulness Justification	Assessment
8) I laugh and joke around with all students in the class.	Distributive	Equality	Interaction
9) I do not care about involving students in syllabus design and implementation.	Interactional	Caring	Teaching
10) When grading, I remove students' names from their exam papers in order not to become influenced by their personal traits.	Procedural	Bias suppression	Assessment
11) If students cheat on exam, I will not disregard and will give notice of the occurrence.	Procedural	Ethicality	Assessment
12) I provide differentiated teaching to students based on their needs to maximize their learning opportunity.	Distributive	Need	Teaching
13) I dedicate the same amount of time to each student for exams.	Distributive	Equality	Assessment
14) I often provide only certain students in the class with access to information on exam materials.	Distributive	Equality	Assessment
15) I grade students based on their achievements.	Distributive	Equity	Assessment
16) After I grade students' exam papers, I give them the opportunity to look at their papers and grieve about their grade.	Procedural	Voice	Assessment
17) I mark students' test papers based on their pe1f01mance on the test.	Distributive	Equity	Assessment
18) I do not treat students differently based on their gender, race, religion, culture, or appearance.	Procedural	Bias suppression	Interaction
19) When I make modifications in my teaching approach, I explicate the logic behind the changes.	Interactional	Justification/ Adequacy	Teaching
20) When a student fails to answer my question in the class, I do not embarrass him/her in front of others.	Interactional	Respect	Interaction
21) The exams in my class include enough explanations and instructions for students to show what they know and what they can do.	Interactional	Justification/ Adequacy	Assessment
22) The allocation of punishment in my class depends on the severity and controllability of students' misbehavior.	Distributive	Equity	Interaction
23) I try my best to provide answers to the questions of all students.	Distributive	Equality	Interaction

24) I make clear what the students should do when they intend to ask me questions.	Procedural	Transparency	Teaching
25) I specify my expectations, syllabus, and grading criteria at the beginning of the semester.	Interactional	Timeliness	Teaching
26) I provide equal attention and help to both high- and low-achievers in my class.	Distributive	Equality	Teaching
27) I have a caring and supportive relationship with my students.	Interactional	Caring	Interaction
28) I announce the content of exams on time.	Interactional	Timeliness	Assessment
29) I am sensitive to my students' feelings, opinions, and rights.	Interactional	Caring	Interaction
30) I provide students in the class with equal access to information on exam materials.	Distributive	Equality	Assessment
31) I adequately inform students of my class attendance policy.	Interactional	Adequacy	Learning
32) I apologize if it is turned out that I have made a mistake or behaved rudely toward a student.	Procedural	Correctability	Interaction
33) If I teach a topic incorrectly in the class, I will accept and attempt to provide correct information.	Procedural	Correctability	Teaching
34) The grading procedures that I use in my classes comply with moral and ethical standards.	Procedural	Ethicality	Assessment
35) The meaning and intention of the feedback that I provide to students are clear.	Procedural	Transparency	Learning
36) I always set equal opportunity and time to students to participate in classroom discussions.	Procedural	Consistency	Learning
37) I have some favorite student that I treat differently from others in class.	Interactional	Propriety	Interaction
38) The test items that I develop match with what has been covered in the course.	Procedural	Consistency	Assessment
39) I set reasonable expectations for my students' performance.	Procedural	Reasonableness	Teaching
40) I give my students the opportunity to express their views and feelings regarding the scheduling of homework and topics I plan to set in the course.	Procedural	Voice	Learning
41) I provide my students with voice in syllabus design and implementation.	Procedural	Voice	Teaching
42) I do not have a favorite student who I treat differently from other students in class.	Distributive	Equality	Interaction
43) If I found that there is no agreement between the class content and exam content, I will accept and remove from exam those questions that I have not taught in the class.	Procedural	Correctability	Assessment
44) I explain everything clearly, so students understand.	Procedural	Transparency	Teaching
45) I allow my students to express their concerns regarding the attendance policy.	Procedural	Voice	Teaching
46) I praise students based on their performance and effort.	Distributive	Equity	Interaction