Tabaran Institute of Higher Education
International Journal of Language Testing
Vol. 11, No. 1, March 2021
ISSN 2476-5880

Assessment Perceptions and Practices in Academic Domain: The Design and Validation of an Assessment Identity Questionnaire (TAIQ) for EFL Teachers

Masoomeh Estaji[1]*, Farhad Ghiasvand[2]

Received: 7 December 2020 Accepted: 27 February 2021

Abstract

The present research aimed to conceptualize the construct of Teacher Assessment Identity (TAI) by designing and validating a questionnaire in the Iranian EFL context. In so doing, a tentative scale with 96 items was piloted on 340 novice and experienced Iranian EFL teachers using Exploratory and Confirmatory Factor Analysis (EFA, CFA). The results of the analyses led to the removal of 33 items, leaving the questionnaire with 61 items on a five-point Likert scale. Moreover, the results revealed that the construct of TAI has 12 factors including assessment "knowledge", "beliefs", "attitudes", "skills and confidence", "practices", "use assurance", "feedback", "rubric/criteria", "consistency and consequence", "grading/scoring", "question-types", and "roles". Likewise, the convergent validity and reliability of the instrument to measure the construct of concern was statistically confirmed ($p>.05$). The findings have various implications for EFL teachers, teacher trainers, course designers, and language researchers by raising their awareness of assessment identity and its underlying components.

Keywords: Assessment identity, EFL teachers, questionnaire design, teacher identity, validation

1. Introduction

The concept of teacher identity has recently positioned itself as an emerging new sub-field of identity theory (Beijaard, Meijer, & Verloop, 2004; Beijaard, Verloop, & Vermunt, 2000; Sachs, 2001). It has become "the bread and butter of our educational diet" (Hoffman, 1998, p. 324). By definition, it refers to "how teachers define themselves to themselves and to others" (Lasky, 2005, p. 901). The construction of this personal sense of the world along with the development of a clear understanding of how teachers see themselves as interacting with others has made teacher identity the core of the teaching profession (Thomas & Beauchamp, 2007). Similarly, Palmer (2007) regards this sense of identity as the trait common to all good teachers. To highlight its significance, he further went on to argue that we are not able to know our students until we know ourselves.

[1] Associate Professor of Applied Linguistics, Department of English Language and Literature, Allameh Tabataba'i University, Tehran, Iran.
[2] Department of English Language and Literature, Allameh Tabataba'i University, Tehran, Iran
* (corresponding author: mestaji74@gmail.com)

This centrality of teacher identity in education has inspired a growing surge of interest in this research domain over the past couple of decades (Beauchamp & Thomas, 2009; Beijaard et al., 2004; Clarke, 2008). Tracing the available literature, one can easily notice that a vast majority of the conducted studies has focused on teacher's professional identity (PI) formation, characteristics, representations, conflicts, and crisis (Beauchamp & Thomas, 2009; Yuan, 2019). What seems to be missing in this research strand is the investigation of Teacher's Assessment Identity (hereafter, TAI) that the teachers craft in relation to the assessment practices that they carry out in the classes. Undoubtedly, teachers as professionals construct and reconstruct many identities in relation to the context within which they work (Hall, 2000). In the context of testing and assessment, too, EFL teachers with varying experiences perceive different identities for themselves as assessors of the students' competence (Adie, 2013). From the perspective of Looney (2014), what is valued or significant for the students to learn is represented through the ways they are assessed.

Nevertheless, most of the studies conducted on teacher identity have exclusively focused on the stages that language teachers undergo for their PI construction and reconstruction (Hevern, 2004; Hramiak, Boultonb, & Irwinc, 2009). Yet, investigating TAI as a crucial dimension of teacher's perceived sense of identity has been overlooked despite the fact that assessment itself is a part of one's PI (Norcini & Burch, 2007; Wood, 2016). In the same manner, developing a multidimensional scale to measure EFL teachers' identity in the realm of assessment and specify the underlying components of this identity-type has been limitedly researched. Considering these shortcomings, the present study was an attempt to design and validate a TAI questionnaire for teachers in the Iranian EFL context.

2. Literature Review

In its very nature, the notion of teacher identity is associated with the conception of "who am I?" and "who are you?" and comprises internal and external realities (Castañeda, 2011). The former denotes cognition, whereas the latter reflects roles. The internal images created by teachers, which are considered as critical in the development of teacher identity, emanate from a prolonged process, which begins in their experiences as being learners (Borg, 2017; Lortie, 1975; Malderez, Hobson, Tracey, & Kerr, 2007). Being an instructor is fundamentally revealed by what encompasses the visible and invisible grounds of work and life (Castañeda, 2011). While the former includes what teachers do (e.g., classroom interaction, evaluation, materials designing, or task implementation), the latter involves more personal criteria (e.g., cognition, beliefs, expectations, or feelings) (Allwright & Bailey, 1991).

Drawing on Britzman (1991), Clarke (2008, p. 8) distinguished between two interpretations of language teacher identity namely, "being the teacher" that is acquiring the skills and competencies to perform the functions of an instructor and "becoming a teacher" that is growing a sense of oneself as an instructor. By such interpretations, Clarke (2008) was of the opinion that learning to teach means "becoming" an instructor rather than acquiring techniques and skills.

ISSN 2476-5880

Generally, it can be contended that teacher identity is "relational, negotiated, constructed, enacted, transforming and transitional" (Miller, 2008, p. 174). It is by no means a fixed property of a teacher but instead a process that evolves, changes, or resists as teachers gain more and more experience, consolidate professional knowledge, and take on plans of personal and professional development (Tsui, 2005).

Teacher identity reveals, "who we are"; as our teaching practice "holds a mirror to the soul" (Palmer, 2007, p. 3). Basically, we shape and reshape our identity in interaction with others in a professional context (Beauchamp & Thomas, 2009). This echoes teacher's PI, which has dominated this research strand for decades (Beijaard et al., 2004). As claimed by many scholars in this area, teacher identity is affected by many factors including context, perception, age, experience, socio-political background, and others (Canrinus, 2011; Mansaray, 2011). If teaching and assessment are two sides of the same coin (Heaton, 1988) and this is the assessment that connects all aspects of the teaching and learning cycle (Brown, 2004), what changes our identity may undergo as assessors of the language. This tough nut has recently been cracked by the introduction of a new concept called Teacher Assessment Identity (TAI) in the field of education.

For a long time, assessment literacy has been utilized to denote teachers' knowledge, understanding, and skills in the context of language assessment (DeLuca & Braund, 2019; Watmani, Asadollahfam, & Behin, 2020). Nevertheless, a number of scholars have extended the concept beyond what it previously denoted (e.g., Khadijeh & Amir, 2015; Siegel & Wissehr, 2011). The offshoot of such a development was the formation of tension between the concept of teacher assessment literacy, which has had an excessively narrow and instrumentalist view, and a broader and more complex notion known as "TAI" (Looney, Cumming, van Der Kleij, & Harris, 2017).

As put by Adie (2013), TAI refers to the perceptions teachers had of themselves as assessors and their concerns about how they might be personally judged by others in the discussion of how they had executed professional judgment (of students' work). It denotes a teacher's assessment knowledge, perceptions, beliefs, practices, confidence, and how he/she is perceived by others as a professional assessor of language. According to Looney et al. (2017), the two concepts of teacher self-efficacy and dispositions are very significant in the conceptualization of TAI. Self-efficacy of teachers refers to the individual's beliefs in his/her abilities to accomplish a specific teaching task with a particular level of quality in a particular context (Dellinger, Bobbett, Olivier, & Ellett, 2008). However, teacher dispositions are the attitudes, perceptions, and beliefs, which form the foundation of one's behavior (the National Council for Accreditation of Teacher Education, NCATE, 2001).

This conceptualization of teachers as assessors represents a broader dimensionalization of teachers' work in assessment in many different aspects. For instance, it is contended that teacher assessment knowledge, confidence, personal disposition, and emotional engagement with assessment influence teachers' own assessment practices in classrooms. This ignited the attempts to specify the possible dimensions of TAI among the researchers interested in this domain. In their landmark study to provide a unified model for the dimensions of TAI, Looney et al. (2017) consulted the studies done on assessment and assessment literacy (e.g., Bandura,1986; Beijaard et

ISSN 2476-5880

al., 2004; Mockler, 2011; Popham, 2009; Stiggins, 1991, 1995) and proposed a model which had 5 dimensions. They came under the headings of "I know", "I feel", "my role", "I believe", and "I am confident".

Looney et al. (2017) regarded the mentioned dimensions as being related to a teacher subject rather than as abstractions to accentuate their argument that who teachers are in the process of assessment is as significant as what they know and are able to do. Such a contention is said to be at the heart of TAI. Furthermore, in their revolutionary research, they cautioned the readers that these dimensions are not mutually exclusive as they have interrelationships with each other, and that identity is a very complex concept, around which it is almost impossible to draw clear boundaries. The components of such a construct are like nested systems, which grow within each other.

Taking a practical approach, Adie (2013) conducted a seminal study on the development of TAI through participation in online moderation. The results of analyses of the data through a sociocultural lens of becoming suggested that participation in online moderation, while challenging for teachers, could provide opportunities to construct and negotiate an identity as an assessor of student work. Moreover, in their seminal study, Jan-nesar, khodabakhshzadeh, Motallebzadeh, and Khajavi (2021) ran similar research in Iran to examine the components of TAI and offered a new empirical TAI model for Iranian EFL teachers. They found that TAI is a three-dimensional model, consisting of *assessment literacy*, *assessment dispositions*, and *contextual factors*. What these influential studies signify is that this research area is still in its infancy and needs further empirical research in order to be established as an independent and significant domain in identity research. Driven by this dearth of research on TAI and the shortage of a validated and multidimensional scale to unpack its underlying components, as also suggested to future researchers by Looney et al. (2017), the present study aimed to design and validate a questionnaire useful for measuring TAI and its various dimensions in an EFL context.

3. Method

3.1 Participants

The main participants of the present study were 340 Iranian novice (89) and experienced (251) EFL teachers, being selected out of 1100 EFL teacher participants, who were working at different language institutes in different cities in Iran. They were both male (107, 31%) and female (233, 69%) instructors with different academic degrees (e.g., BA, MA, and PhD). They mostly majored in English language teaching (310), Translation (9), Literature (16), and Linguistics (5). Their age ranged from 21 to 57 with a mean of 31.95. It is critical to note that the participants of this study were chosen non-randomly using convenience sampling and based on their willingness to partake in the study.

3.2 Instruments

This study followed two core stages in its procedure. First, a Teacher Assessment Identity Questionnaire (TAIQ) was designed by the researchers, and then it was validated precisely in relation to the collected data from 340 Iranian EFL teachers. The newly designed questionnaire

ISSN 2476-5880

had 12 sections including 61 items on a five-point Likert scale. The succeeding sections present the steps taken in designing and validating the scale in-details.

3.3 Data Collection Procedure

In the initial phase of the study, in order to design a valuable questionnaire in the domain of teacher assessment identity, a substantial review of the literature related to teacher assessment was conducted to get to know the related constructs. The aim of this stage was to examine the current frameworks and models and the research tools that might previously exist in this domain. It is essential to note that in the item generation phase, the researchers took advantage of a blend of deductive and inductive approaches. As put by Cheng (2017), the deductive approach to item generation includes a broad review of the literature, whereas an inductive approach depends on individual answers like requesting a representative sample from the population to describe their feelings or behaviors. Consequently, the researchers not only went deeply through the literature on TAI but also carried out an interview with some EFL teachers and testing experts to achieve more information. It is noteworthy that the standard procedures of developing a valid and reliable research tool proposed by Dörnyei (2003) were carefully followed in this study. To design the first draft of TAI constructs and their related concepts, the literature review, initial interviews with the experts, and the theoretical framework introduced by Looney et al. (2017) facilitated the ground for the researchers and formed the basis for item generation.

Next, the researchers defined the TAI construct both theoretically and operationally. Then the dimensionality of TAI was examined to determine whether or not the constructs could be comprised of several related components. Subsequently, the format and the number of the items of the questionnaire were determined. Having the items written, the researchers omitted the repetitive ones and the list was condensed to 111 items. Afterward, the content validity of the questionnaire was checked by 5 experts in the field of Applied Linguistics who had PhD degrees and were teaching at higher education levels in Tehran universities. In so doing, the first version of the instrument was given to the panel of experts to examine its content and language suitability by rating each item from 1 to 5 and propose their views and comments for each item or component to improve the questionnaire. According to Koellne and Jacobs (2015), if all these steps are systematically completed, the content validity of the research instrument is recognized and proven. After gaining and analyzing the experts' opinions, some items were deleted or revised. Eventually, 96 items remained for being included in the final version of the questionnaire.

The next step was to test the clarity, relevance, and readability of the items of the questionnaire by piloting it on a small representative sample taken from the target population. To do so, 30 EFL teachers were asked to take part in a piloting phase of the instrument. As Creswell (2003) pinpointed, the goal of a pilot study is to increase the quality of the questions and the format of the survey. The online link of the questionnaire was sent to each participant in this phase through email and Telegram application to be answered carefully. Meanwhile, the researchers requested the participants to give their comments about the directions and the length of the instrument as well. Finally, the 96-item questionnaire was distributed among 1100 novice and experienced EFL teachers in different cities of Iran to be completed in their free time. After about two months, 340

ISSN 2476-5880

questionnaires were filled in and submitted to the researchers both in online and printed versions. The questionnaire, then, went through the following statistical analyses for measuring its construct validity.

3.4 Data Analysis

In order to analyze the collected data of this study, EFA and CFA were employed to determine the construct validity and strength of the relationship in the questionnaire items. Through EFA, the underlying factors were explored to reduce the dimensions and extract the relevant factors. Similarly, the precision of the measurement of the structures was examined by the relevant indices. In this phase, CFA specifies whether or not the designed items can actually measure what they claim to measure. Likewise, the relevance of the extracted factors to the variables and factors was determined.

4. Results

4.1 Exploratory Factor Analysis

A 96-item questionnaire was developed to capture the participants' perceptions of teacher assessment identity. The obtained results were analyzed through factor analysis to ensure its construct validity. First, the KMO value of 0.79 > 0.6 showed sample adequacy. The result of Bartlett's Test of sphericity (χ^2 (4560) = 14691.91, p = .00) was also significant. Running factor analysis with Principle Component Analysis (PCA) extracted 28 factors with eigenvalue above 1, explaining 69.35% of the total variances. Consulting with parallel analysis to PCA, the number of factors was reduced to 13 with eigenvalues above 1.69. Running Oblimin rotation, the following correlation matrix was generated (Table 1).

Table 1.
Component Correlation Matrix Based on Oblimin Rotation

Component	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2	.08												
3	.02	.08											
4	-.03	-.03	-.00										
5	-.30	-.16	-.07	.00									
6	.20	.09	.03	.03	-.13								
7	.10	.15	.13	-.04	-.13	.07							
8	.12	.09	.04	-.00	-.13	.12	.10						
9	.04	.26	.14	-.02	-.05	.02	.12	.08					
10	.18	.11	.01	-.02	-.14	.13	.11	.08	-.00				
11	.01	-.06	-.08	.00	.06	.00	-.09	-.09	-.06	.01			
12	-.18	-.23	-.13	-.01	.22	-.15	-.21	-.10	-.15	-.10	.06		
13	.13	.06	.05	.01	-.07	.06	.06	.05	.04	.01	-.01	-.10	

Extraction Method: Principal Component Analysis.
Rotation Method: Oblimin with Kaiser Normalization.

ISSN 2476-5880

As it is evident from Table 1, there was no semi-significant correlation between the pairs of components; therefore, the EFA was proceeded with an orthogonal rotation, i.e., varimax rotation. Appendix A (Table 1) shows the variance explained from the EFA with 13 fixed factors. As reported in Table 2, the 13 components explain 49% of the cumulative variance. The inspection of the scree plot (Figure 1) shows that the four initial components had the highest eigenvalues, with some slight fluctuations in the explained variances from the fifth component to the thirteenth one. The dots then started decreasing steadily. That is why the thirteen factors seem appropriate to be extracted. Note that the thirteenth factor is the last factor, which is over the eigenvalue obtained from the parallel analysis to PCA, i.e., 1.69. Besides, looking into the total variance explained in the initial analysis, the first 4 factors only explained the cumulative average variance of 29.1%, which is far less than enough to be generalizable; therefore, the researchers decided to re-run the analysis with 13 fixed factors.

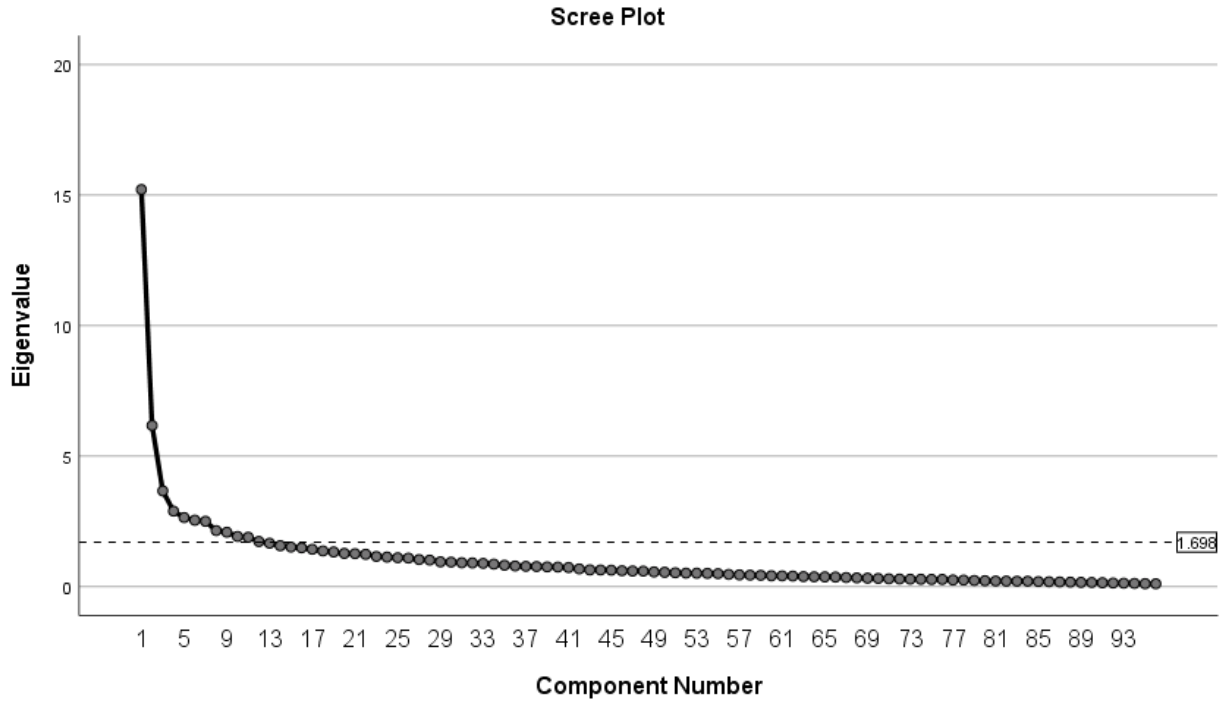

Figure 1. Scree Plot

Finally, in order to explore which item loads up to which factor, the factor loadings were checked out (Appendix B, Table 2). The varimax rotation was done by suppressing the loading coefficients below 0.4. This resulted in discarding 22 items with low loadings. Moreover, two items (items 6 and 8) showed characteristics of complex items loading up with close coefficients to more than one component. These items were also discarded.

ISSN 2476-5880

4.2 Confirmatory Factor Analysis

As mentioned, the initial EFA resulted in the identification of 13 factors and exclusion of 24 items. The remaining items went through a CFA using a covariance-based software, i.e., IBM AMOS. Figure 2 shows the initial model with standardized estimates.

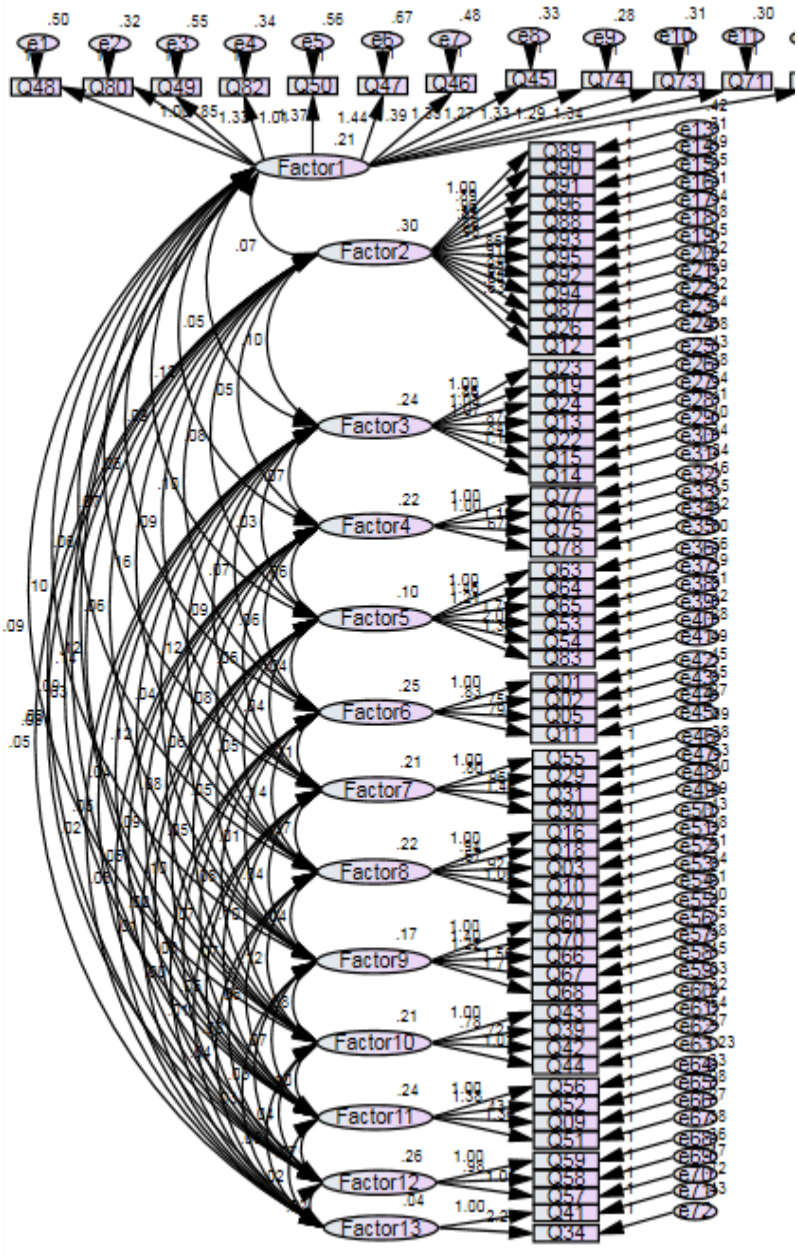

Figure 2. Standardized Estimates of the Initial Model

ISSN 2476-5880

4.3 Evaluation of Convergent Validity

According to Hair, Black, Babin, and Anderson (2009), two conditions must be met for a model to have convergent validity. First, the loaded value must be significant; and second, the loaded value must be larger than 0.5. In the following, the description of factor loadings for each component will be evaluated.

As indicated in Table 2, all items had significant loaded values. The inspection of standardized loadings (standardized estimates) showed no item had loadings less than 0.5. Therefore, all items in this factor were safely kept. Moreover, the R-Squared values showed that item 72 had the highest and item 42 had the lowest shares in the estimation of Factor 1.

Table 2.
Initial Factor Loading for Factor 1

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q48	<---	Factor1	1.00				.54	.29
Q80	<---	Factor1	.84	.10	8.25	.00	.56	.31
Q49	<---	Factor1	1.33	.14	8.92	.00	.63	.40
Q82	<---	Factor1	1.01	.11	8.83	.00	.62	.38
Q50	<---	Factor1	1.37	.15	9.00	.00	.64	.41
Q47	<---	Factor1	1.43	.16	8.85	.00	.62	.39
Q46	<---	Factor1	1.38	.15	9.26	.00	.67	.45
Q45	<---	Factor1	1.35	.13	9.71	.00	.73	.53
Q74	<---	Factor1	1.27	.13	9.76	.00	.73	.54
Q73	<---	Factor1	1.33	.13	9.75	.00	.73	.54
Q71	<---	Factor1	1.29	.13	9.72	.00	.73	.53
Q72	<---	Factor1	1.33	.13	10.05	.00	.77	.60

Table 3 also shows that all items had significant loaded values. The inspection of standardized loadings showed that items 87 and 12 had loadings less than 0.5, thus, were eliminated from the model. Moreover, the R-Squared values showed that item 93 had the highest and item 26 had the lowest shares in the estimation of Factor 2.

Table 3.
Initial Factor Loading for Factor 2

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q89	<---	Factor2	1.00				.64	.41
Q90	<---	Factor2	.89	.08	10.25	.00	.65	.43
Q91	<---	Factor2	.96	.10	9.53	.00	.60	.36
Q96	<---	Factor2	.85	.08	9.77	.00	.61	.38
Q88	<---	Factor2	.85	.08	10.05	.00	.64	.41
Q93	<---	Factor2	.95	.09	10.35	.00	.66	.44

ISSN 2476-5880

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q95	<---	Factor2	.86	.09	9.02	.00	.56	.31
Q92	<---	Factor2	.91	.09	9.51	.00	.59	.35
Q94	<---	Factor2	.83	.08	9.90	.00	.62	.39
Q87	<---	Factor2	.61	.09	6.64	.00	.40	.16
Q26	<---	Factor2	.64	.07	8.50	.00	.52	.27
Q12	<---	Factor2	.62	.09	6.97	.00	.42	.17

As Table 4 shows, all items had significant loadings and none of them had standardized estimates less than 0.5; therefore, all items were kept safely in the model. Moreover, inspection of R-Squared values showed that item 24 had the highest share for estimation of Factor 3 while item 23 had the lowest.

Table 4.
Initial Factor Loading for Factor 3

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q23	<---	Factor3	1.00				.51	.26
Q19	<---	Factor3	.86	.12	7.10	.00	.54	.29
Q24	<---	Factor3	1.02	.13	7.75	.00	.63	.40
Q13	<---	Factor3	1.06	.14	7.15	.00	.54	.30
Q22	<---	Factor3	.86	.11	7.56	.00	.60	.36
Q15	<---	Factor3	.83	.11	7.13	.00	.54	.29
Q14	<---	Factor3	1.09	.13	8.04	.00	.68	.46

Based on the results reported in Table 5, all four items in Factor 4 had had significant loaded values with standardized estimates above 0.5. Moreover, item 75 had the highest R-squared value and item 78 had the lowest.

Table 5.
Initial Factor Loading for Factor 4

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q77	<---	Factor4	1.00				.69	0.48
Q76	<---	Factor4	1.00	.08	11.77	***	.76	0.58
Q75	<---	Factor4	1.12	.09	12.10	***	.80	0.64
Q78	<---	Factor4	.66	.08	7.99	***	.51	0.26

ISSN 2476-5880

The results reported in Table 6 shows that all items had significant loaded values. However, items 63 and 65 did not meet the criterion of at least 0.5 standardized estimation; therefore, they were excluded from the model. Moreover, the highest R-Squared value belonged to item 54 and the lowest to item 64.

Table 6.
Initial Factor Loading for Factor 5

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q63	<---	Factor5	1.00				.33	.11
Q64	<---	Factor5	1.46	.28	5.12	.00	.53	.28
Q65	<---	Factor5	1.21	.24	4.96	.00	.48	.23
Q53	<---	Factor5	1.70	.32	5.25	.00	.57	.33
Q54	<---	Factor5	2.04	.36	5.59	.00	.75	.57
Q83	<---	Factor5	1.29	.24	5.35	.00	.61	.38

As Table 7 indicates, all items had significant loaded values with standardized estimates above 0.5. Besides, items 11 and 2 had the highest and lowest R^2 values, respectively.

Table 7.
Initial Factor Loading for Factor 6

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q01	<---	Factor6	1.00				.57	.33
Q02	<---	Factor6	.82	.12	6.70	.00	.52	.27
Q05	<---	Factor6	.74	.11	6.77	.00	.53	.28
Q11	<---	Factor6	.79	.10	7.29	.00	.60	.36

As illustrated in Table 8, all items had significant loaded values. However, item 55 was excluded as it had standardized estimate value below 0.5. Besides, items 30 and 31 had the highest and lowest R^2 values, respectively.

Table 8.
Initial Factor Loading for Factor 7

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q55	<---	Factor7	1.00				.43	.19
Q29	<---	Factor7	.80	.14	5.75	.00	.51	.26
Q31	<---	Factor7	.95	.17	5.58	.00	.50	.23
Q30	<---	Factor7	1.40	.21	6.41	.00	.76	.58

ISSN 2476-5880

As Table 9 demonstrates, all items had significant loaded values. However, item 3 was excluded as it had standardized estimate value below 0.5. Besides, items 20 and 10 had the highest and lowest R^2 values, respectively.

Table 9.
Initial Factor Loading for Factor 8

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q16	<---	Factor8	1.00				.55	.30
Q18	<---	Factor8	.93	.12	7.32	.00	.55	.30
Q03	<---	Factor8	.66	.14	4.55	.00	.30	.09
Q10	<---	Factor8	.92	.13	6.98	.00	.51	.26
Q20	<---	Factor8	1.00	.12	7.87	.00	.62	.38

The results reported in Table 10 shows that all items had significant loaded values. However, item 60 did not meet the criterion of at least 0.5 standardized estimation; therefore, it was excluded from the model. Moreover, the highest R-Squared value belonged to item 68 and the lowest to item 66.

Table 10.
Initial Factor Loading for Factor 9

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q60	<---	Factor9	1.00				.49	.24
Q70	<---	Factor9	1.40	.20	6.71	.00	.52	.27
Q66	<---	Factor9	1.22	.18	6.56	.00	.50	.25
Q67	<---	Factor9	1.56	.19	7.90	.00	.72	.52
Q68	<---	Factor9	1.76	.22	7.95	.00	.73	.54

As shown in Table 11, all items had significant loaded values with only item 42 having standardized estimate below 0.5. Moreover, inspection of R-Squared values showed that item 44 had the highest share for estimation of Factor 10 while item 39 had the lowest.

Table 11.
Initial Factor Loading for Factor 10

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q43	<---	Factor10	1.00				.62	.38
Q39	<---	Factor10	.78	.10	7.47	.00	.53	.28
Q42	<---	Factor10	.71	.11	6.03	.00	.40	.16
Q44	<---	Factor10	1.00	.11	8.62	.00	.66	.44

ISSN 2476-5880

Based on the results reported in Table 12, all four items in Factor 4 had significant loaded values. Item 9 was excluded as it had standardized estimate value below 0.5. Moreover, item 51 had the highest R-squared value and item 56 had the lowest.

Table 12.
Initial Factor Loading for Factor 11

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q56	<---	Factor11	1.00				.50	.25
Q52	<---	Factor11	1.38	.20	6.62	.00	.76	.58
Q09	<---	Factor11	.42	.13	3.14	.00	.20	.04
Q51	<---	Factor11	1.35	.20	6.65	.00	.78	.61

As indicated in Table 13, all items had significant loading values with standardized estimates above 0.5. Besides, items 58 and 59 similarly had the highest R^2 values, while item 57 had the lowest R^2 values respectively.

Table 13.
Initial Factor Loading for Factor 12

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q59	<---	Factor12	1.00				.63	.40
Q58	<---	Factor12	.98	.14	6.86	.00	.63	.40
Q57	<---	Factor12	1.01	.14	6.83	.00	.59	.35

Table 14 shows both items with significant loaded values. However, item 41 was excluded as it had standardized estimate value below 0.5. Consequently, this factor was left with only one item. As one item cannot be considered as a factor, the researcher discarded the whole factor from the model.

Table 14.
Initial Factor Loading for Factor 13

Observed		Latent Variable	Unstandardized Estimates				Standardized Estimates	
			Estimate	S.E.	T	P	Estimate	R Squared
Q41	<---	Factor13	1.00				.22	.04
Q34	<---	Factor13	2.20	.82	2.65	.00	.54	.29

ISSN 2476-5880

4.4 Evaluation of Discriminant Validity

In contrast with convergent validity, discriminant validity is the extent to which a test can differentiate between unrelated traits (Hayashi & Hays, 1987). It can be evaluated by using Fornell and Larcker's (1981) criterion, which compares the square root of the average variance extracted (AVE) with the correlation of latent constructs (Table 15).

Table 15.
Testing the Discriminant Validity Against Fornell and Larcker's Criterion

	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7	Factor 8	Factor 9	Factor 10	Factor 11	Factor 12
Factor1	0.65											
Factor2	0.20	0.58										
Factor3	0.56	0.30	0.69									
Factor4	0.58	0.17	0.39	0.62								
Factor5	0.30	0.31	0.23	0.32	0.50							
Factor6	0.27	0.44	0.28	0.21	0.46	0.59						
Factor7	0.27	0.52	0.40	0.39	0.40	0.36	0.53					
Factor8	0.51	0.20	0.31	0.27	-0.00	0.15	0.24	0.63				
Factor9	0.47	0.55	0.40	0.36	0.33	0.52	0.51	0.35	0.60			
Factor10	0.52	0.15	0.39	0.54	0.31	0.23	0.28	0.30	0.45	0.67		
Factor11	0.13	0.17	0.20	0.07	0.05	0.25	0.12	0.15	0.17	0.26	0.62	
Factor12	0.27	0.36	0.20	0.46	0.49	0.38	0.50	0.23	0.44	0.32	0.10	0.60

As reported in Table 15, none of the factors had correlations above the squared root of AVE (bold value in the table) with another factor. Therefore, the criterion for the discriminant validity (Henseler, Ringle, & Sarstedt, 2015) was also met.

4.5 Modification of the Model

As reported above, 10 items did not meet the two criteria for convergent validity; hence discarded from the model. Besides, the only remaining item in Factor 13 was excluded, leaving the model with 12 factors and 61 items. Furthermore, using the suggestions made by the software suggested modifications were done. Figure 3 shows unstandardized estimates (variances explained) of the model after modifications are applied.

ISSN 2476-5880

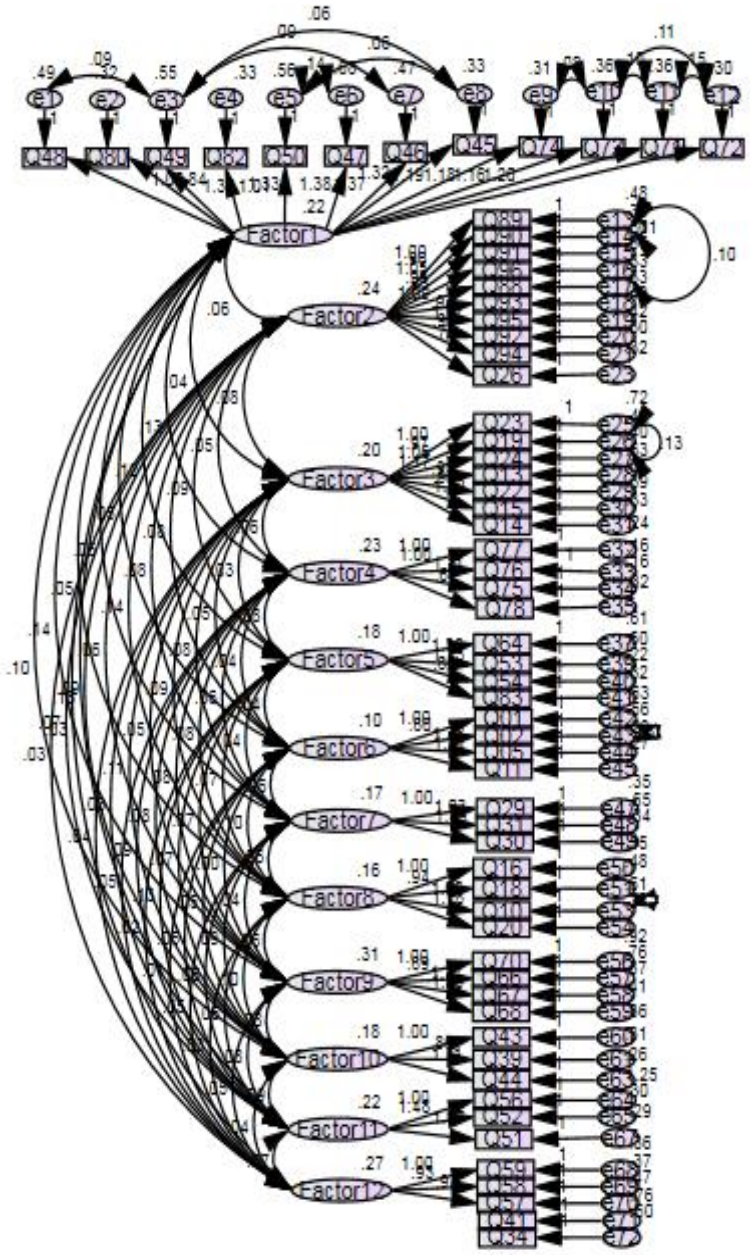

Figure 3. Unstandardized Estimates of the Modified Model

The variances estimated by each component are also presented in Table 15.

ISSN 2476-5880

Table 16.
Variances Explained by Each Factor (After Modification)

	Estimate	S.E.	P	R^2	Cumulative Variance
Factor1	.22	.04	.00	0.04	0.04
Factor2	.24	.04	.00	0.05	0.10
Factor3	.20	.05	.00	0.04	0.14
Factor4	.23	.03	.00	0.05	0.19
Factor5	.18	.04	.00	0.03	0.22
Factor6	.10	.03	.00	0.01	0.23
Factor7	.17	.03	.00	0.02	0.26
Factor8	.16	.04	.00	0.02	0.29
Factor9	.31	.07	.00	0.09	0.39
Factor10	.18	.03	.00	0.03	0.42
Factor11	.22	.06	.00	0.04	0.47
Factor12	.27	.05	.00	0.07	0.54

As reported in Table 16, the modified model cumulatively explains 54.84% of the variance in the data.

4.5 Goodness of Fit

The three essential criteria for model fit are reported in Table 17. As reported, all three criteria are met, indicating the goodness of fit for the model.

Table 17.
Evaluation of the Model's Goodness of Fit

Criteria	Observed Value	Required Value	Evaluation
CMIN/df	2.03	< 3	Met
RMSEA	.05	< .08	Met
PNFI	.58	> .5	Met

4.6 Model Reliability

The model reliability was measured by the Composite Reliability (CR) for each factor (Table 18). As reported, all factors showed acceptable indices of reliability.

Table 18.
Composite Reliability of the Model

	Factor 1	Factor2	Factor3	Factor4	Factor5	Factor 6	Factor7	Factor8	Factor9	Factor10	Factor11	Factor12
CR	0.90	0.85	0.77	0.78	0.71	0.65	0.61	0.60	0.72	0.63	0.69	0.65

ISSN 2476-5880

Generally, in social sciences, a reliability index of 0.60 to 0.70 is considered acceptable and 0.80 or greater is regarded as a very good reliability level. However, values greater than 0.95 are not essentially good, as they might be a sign of redundancy (Ghazali, 2008).

5. Discussion

The present study was a bid to design and validate a TAI questionnaire, unpacking the underlying components of TAI, which has been, to date, overlooked in the extant literature. This empirical study was administered in the EFL context of Iran, using EFA, CFA, and SEM to analyze the data. It mainly tried to bridge the yawning gap in teacher identity research especially in a close relation to assessment practices by proposing and validating a brand-new TAI questionnaire. Exploratory Factor Analysis was run to detect the core factors and condense the data. Before using the data for factor analysis, the Bartlett and KMO tests were applied to check and approve the adequacy of the collected data. The results of Bartlett's Test of sphericity as well as the KMO were significant and indicative of sample adequacy. Then through Principle Component Analysis (PCA), 28 factors were extracted with eigenvalues above 1, explaining 69.35% of the total variances. Consulting with parallel analysis to PCA, the number of factors were reduced to 13 with eigenvalues above 1.69.

Having the EFA ended up with 13 factors and the exclusion of 24 items, the remaining items went through a CFA by SEM using a covariance-based software (i.e., IBM SPSS AMOS 21). The results indicated that 10 items did not meet the two criteria for convergent validity; hence discarded from the model. Besides, the only remaining item in Factor 13 was excluded, leaving the model with 12 factors and 61 items. After model modifications, it was found that the extracted factors and the remained items were both valid and reliable. Moreover, the fitting indices of the model revealed that the measurement models were confirmed. In other words, the model enjoyed external validity.

To be more specific, the 12 components or factors of the instrument in this study included the assessment "knowledge", "beliefs", "attitudes", "skills and confidence", "practices", "use assurance", "feedback", "rubric/criteria", "consistency and consequence", "grading/scoring", "question-types", and "roles". The results are consistent with those of Looney et al. (2017), who carried out a groundbreaking study on the reconceptualization of TAI as a concept beyond assessment literacy. In their study, they reviewed different investigations on assessment and assessment literacy (e.g., Bandura, 1986; Beijaard et al., 2004; Mockler, 2011; Popham, 2009; Stiggins, 1991, 1995) and came up with a 5-dimensional model of TAI whose dimensions have some overlaps with the extracted factors of the current study. Their model considered "I know", "I feel", "my role", "I believe", and "I am confident" as the main dimensions of TAI without proposing a scale to measure the construct. Nevertheless, these dimensions are comparable to the extracted factors of "knowledge", "beliefs", "attitudes", "confidence", and "role" in this study. The results are also partly in tune with those of Jan-nesar, khodabakhshzadeh, Motallebzadeh, and Khajavi (2021), who carried out a similar study in the same context of the present research and identified three dimensions of TAI including "assessment literacy, dispositions, and contextual factors. However, the present study found 12 components for TAI, and unlike Jan-nesar,

khodabakhshzadeh, Motallebzadeh, and Khajavi (2021) who regarded assessment literacy as one of the three components of their scale, the current study considered TAI as a concept beyond assessment literacy which is a very broad concept that includes many sub-components like assessment knowledge, skills, practices, and roles among others. This conceptualization is supported by Looney et al. (2017), who argued that TAI is a multidimensional construct beyond assessment literacy.

The results of this study are fresh in that they added some other components to the complex construct of TAI, such as assessment "skills", "practices", "use assurance", "feedback", "rubric/criteria", "consistency and consequence", "grading/scoring", and "question-types", and regarded them as very important in shaping and reshaping one's identity as an assessor. A possible justification for the findings of this study can be the fact that Looney et al. (2017) only reviewed the existing studies on assessment literacy and carried out a review study without asking the teachers to express their perceptions of identity as assessors. Therefore, their model did not cover all aspects of TAI. Another reason for the findings of this study might be the EFL context of Iran in which the participants craft their TAI through different assessment practices. This is well substantiated in the literature in that the identity formation process varies from a context to another context (Gibson, Dollarhide, & Moss, 2010).

Looking at the extracted factors in this study, it can be inferred that most of the new factors or components are concerned with the practical side of assessment in the classroom. For instance, EFL teachers' use of oral or written feedback, rubric, different question-types, and the assurance of doing so form the foundation of Iranian EFL teachers' perceived sense of identity as language assessors. This can be attributed to the external reflection of identity in one's actions. Aside from the internal realities of being an assessor, Iranian EFL teachers specifically carved their assessment identity by being engaged in different assessment practices that form a teacher's identity as an assessor.

Undoubtedly, our identity is reflected in our beliefs, talks, and actions in different contexts. Hence, performing different assessment practices can affect the construction or reconstruction of TAI. It is widely accepted that those reputable teachers, who are competent in different assessment areas, use appropriate assessment feedback in the class, have self-made assessment rubrics, vary their assessment questions, are certain about their assessment choices, and know the consequences of their assessments especially in the context of high-stakes examinations. Hence, one's use of unique tips and tricks to prepare the students for a high-stakes test leads to the development of PI, and that the use of different assessment practices can promote one's assessment identity.

Another important finding in this study was the inclusion of the emotional aspects of identity in assessment, which was represented through the "attitude" component. This echoes Beijaard et al.'s (2004) contention that the emotional dimensions of identity are very significant in any analysis or discussion of teacher identity. A place should always be given to the issue of how one "feels" to be a teacher or assessor at any point in time. Admitting this, the items remained in the "attitude" component just reflect Iranian EFL teachers' attitudes or feelings about alternative assessment. This is attributed to their care for the use of novel and alternative forms of assessment instead of the traditional ones. Hence, most of the participants might have selected those items that

were measuring the teacher's attitudes toward alternative assessment, which led to the removal of all the other items under this component.

Another construct that was extracted in this study was the teacher's assessment "confidence", which is comparable to the concept of teacher self-efficacy or confidence in one's ability to perform an assessment task effectively. This is closely interconnected with the other dimensions of identity as one's assessment beliefs, knowledge, and attitudes form the confidence required for the assessment practices. Therefore, our beliefs and knowledge of assessment are of importance as long as they end up in assurance of what we are doing in our assessment.

All in all, the extracted components indicated that TAI has been considered an important part of one's PI. In particular, it considers teachers' identity not merely in relation to what they do, but who they are as language assessors.

6. Conclusion and Implications

In light of the present study, which was a bid to explore the underlying components of TAI in the EFL context of Iran, it can be concluded that the construct of identity, in general, and assessment identity, in particular, involve complex and nested sub-components. Hence, TAI is constructed and developed in relation to many other components such as assessment knowledge, beliefs, attitudes, confidence, actual practices, and roles in the class. These components have strong interrelationships with each other as well. They can be seen as the steps of a ladder, which ultimately end in TAI as an overarching construct in ELT.

This study, as the first empirical attempt to design and validate a TAI questionnaire, has both theoretical and practical implications for the field. Theoretically, the findings of this study can add fresh insights to the growing body of research on teacher identity and identity theory as a whole. Moreover, it sparks a light in academia that in the same way as teachers' pedagogical practices in the class shape their PI, their assessment choices and dispositions lead to the construction of a different identity among them as language assessors. Likewise, this study can expand the literature related to teacher assessment literacy by connecting it to teacher assessment identity research. Avid researchers are invited to pay special attention to assessment identity, which has been remained under the shadow of PI to date.

Practically, the results of this study are of use for EFL teachers, teacher educators, teacher development (TD) course designers, and educational researchers. Concerning the teachers, the results can raise their awareness of assessment identity and its underlying components. Teacher educators can also use the results to identify the trainees' assessment identity, cover different components of TAI, and teach them how to use assessment to promote their identity as assessors. Similarly, TD course designers can inject the components of TAI in their programs to develop this type of identity along with PI. Finally, teacher identity researchers can use the designed instrument to explore English language teachers' perceived level of assessment identity. They can run similar studies with ESL or ESP participants or use qualitative data collection methods to construct a more vivid and reliable picture of TAI in various contexts.

 ISSN 2476-5880

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Adie, L. (2013). The development of teacher assessment identity through participation in online moderation. *Assessment in Education: Principles, Policy, & Practice, 20*(1), 91-106. https://doi.org/10.1080/0969594x.2011.650150

Allwright, R., & Bailey, K. M. (1991). *Focus on the language classroom: An introduction to classroom research for language teachers.* Cambridge: Cambridge University Press.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory.* Upper Saddle River, NJ: Prentice-Hall Inc.

Beauchamp, C. & Thomas, L. (2009, July). *Imagination and reflection in teacher education: The development of professional identity from student teaching to beginning practice*. Symposium Proceedings of the 4th International Conference on Imagination and Education, Opening Doors to Imaginative Education: Connecting Theory to Practice. Vancouver: British Columbia.

Beijaard, D., Meijer, P. C., & Verloop, N. (2004). Reconsidering research on teachers' professional identity. *Teaching and Teacher Education, 20*(2), 107-128. https://doi.org/10.1016/j.tate.2003.07.001

Beijaard, D., Verloop, N., & Vermunt, J. D. (2000). Teachers' perceptions of professional identity: An exploratory study from a personal knowledge perspective. *Teaching and teacher education, 16*(7), 749-764. https://doi.org/10.1016/s0742-051x(00)00023-8

Borg, S. (2017). Identity and teacher research. In G. Barkhuizen (Ed.), *Reflections on language teacher identity research* (pp. 126-132). London: Routledge.

Britzman, D. P. (1991). Decentering discourses in teacher education: Or, the unleashing of unpopular things. *Journal of Education, 173*(3), 60-80. https://doi.org/10.1177/002205749117300305

Brown, G. T. (2004). Teachers' conceptions of assessment: Implications for policy and professional development. *Assessment in Education: Principles, Policy & Practice, 11*(3), 301-318. https://doi.org/10.1080/0969594042000304609

Canrinus, E. T. (2011). *Teachers' sense of their professional identity.* Groningen: PrintPartners Ipskamp.

Castañeda, J. A. F. (2011). *Teacher identity construction: Exploring the nature of becoming a primary school language teacher* (Unpublished Doctoral dissertation). University of Newcastle upon Tyne, United Kingdom.

Cheng, Y. S. (2017). Development and preliminary validation of four brief measures of L2 language-skill-specific anxiety. *System, 68*(1), 15-25. https://doi.org/10.1016/j.system.2017.06.009

ISSN 2476-5880

Clarke, M. (2008). *Language teacher identities: Co-constructing discourse and community*. UK: Multilingual Matters.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches.* Thousand Oaks, CA: Sage.

Dellinger, A. B., Bobbett, J. J., Olivier, D. F., & Ellett, C. D. (2008). Measuring teachers' self-efficacy beliefs: Development and use of the TEBS-Self. *Teaching and Teacher Education, 24*(3), 751-766. https://doi.org/10.1016/j.tate.2007.02.010

DeLuca, C., & Braund, H. (2019). Preparing assessment literate teachers. *Oxford Research Encyclopedia of Education.* https://doi.org/10.1093/acrefore/9780190264093.013.487

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

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/10.2307/3151312

Gibson, D. M., Dollarhide, C. T., & Moss, J. M. (2010). Professional identity development: A grounded theory of transformational tasks of new counselors. *Counselor Education and Supervision, 50*(1), 21-38. https://doi.org/10.1002/j.1556-6978.2010.tb00106.x

Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2009). *Multivariate data analysis.* Upper Saddle River, NJ: Prentice hall.

Hall, S. (2000). Who needs identity? In P. Du Gay, G. Evans, & P. Redman (Eds.), *Identity: A Reader* (pp. 15-30). London: Sage and the Open University.

Hayashi, T., & Hays, R. D. (1987). A microcomputer program for analyzing multitrait-multimethod matrices. *Behavior Research Methods, 19*(3), 345-348. https://doi.org/10.3758/BF03202573

Heaton, J. B. (1988). *Writing English Language Tests.* Harlow: Longman.

Henseler, J., Ringle, C. M., & Sarstedt M. (2015). A new criterion for assessing discriminant validity in variance-based structural equation modeling. *Journal of the Academy of Marketing Science, 43*(1), 115-135. https://doi.org/ 10.1007/s11747-014-0403-8

Hevern, V. W. (2004). Threaded identity in cyberspace: Weblogs & positioning in the dialogical self. *Identity, 4*(4), 321-335. https://doi.org/10.1207/s1532706xid0404_2

Hoffman, D. M. (1998). A therapeutic moment? Identity, self, and culture in the anthropology of education. *Anthropology and Education Quarterly, 29*(3), 324-346. https://doi.org/10.1525/aeq.1998.29.3.324

Hramiak, A., Boulton, H., & Irwin, B. (2009). Trainee teachers' use of blogs as private reflections for professional development. *Learning, Media, and Technology, 34*(3), 259-269. https://doi.org/10.1080/17439880903141521

Jan-nesar, M. Q., Khodabakhshzadeh, H., Motallebzadeh, K. & Khajavy G. (2021). Measuring EFL teachers' assessment identity: development and validation of a questionnaire. *Journal of Language and Translation, 11*(1), 29-46.

Khadijeh, B., & Amir, R. (2015). Importance of teachers' assessment literacy. *International Journal of English Language Education, 3*(1), 139-146.

ISSN 2476-5880

https://doi.org/10.5296/ijele.v3i1.6887

Koellner, K., & Jacobs, J. (2015). Distinguishing models of professional development: the case of an adaptive model's impact on teachers' knowledge, instruction, and student achievement. *Journal of Teacher Education, 66*(1) 51-67. https://doi.org/10.1177/0022487114549599

Lasky, S. (2005). A sociocultural approach to understanding teacher identity, agency and professional vulnerability in a context of secondary school reform. *Teaching and teacher education, 21*(8), 899-916. https://doi.org/10.1016/j.tate.2005.06.003

Looney, A. (2014). Assessment and the reform of education systems. In C. Wyatt-Smith, V. Klenowski, & P. Colbert (Eds.), *Designing assessment for quality learning* (pp. 233-247). Heidelberg: Springer.

Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2017). Reconceptualising the role of teachers as assessors: Teacher assessment identity. *Assessment in Education: Principles, Policy & Practice, 25*(5), 442-467. https://doi.org/10.1080/0969594x.2016.1268090

Lortie, D. C. (1975). *School teacher.* Chicago, IL, University of Chicago Press.

Malderez, A., Hobson, A. J., Tracey, L., & Kerr, K. (2007). Becoming a student teacher: Core features of the experience. *European Journal of Teacher Education, 30*(3), 225-248. https://doi.org/10.1080/02619760701486068

Mansaray, A. (2011). *Teacher Identity: An African perspective.* A paper presented at an International workshop on dialogue for educational development at Hiroshima University, Japan.

Miller, L. (2008). Developing professionalism within a regulatory framework in England: challenges and possibilities. *European early childhood education research journal, 16*(2), 255-268. https://doi.org/10.1080/13502930802141667

Mockler, N. (2011). Beyond what works: Understanding teacher identity as a practical and political tool. *Teachers and Teaching, 17*(5), 517-528. https://doi.org/10.1080/13540602.2011.602059

National Council for Accreditation of Teacher Education (NCATE). (2001). Standards for professional development schools. Washington, DC: NCATE. Retrieved from http://www.ncate.org/documents/pdsstandards.pdf

Norcini, J., & Burch, V. (2007). Workplace-based assessment as an educational tool: AMEE Guide No. 31. *Medical Teacher, 29*(9-10), 855-871. https://doi.org/10.1080/01421590701775453

Palmer, P. J. (2007). *The courage to teach: Exploring the inner landscape of a teacher's life.* San Francisco: Jossey-Bass.

Popham, J. (2009). Assessment literacy for teachers: Faddish or fundamental? *Theory into Practice, 48*(1), 4-11. https://doi.org/10.1080/00405840802577536

Sachs, J. (2001). Teacher professional identity: Competing discourses, competing outcomes. *Journal of Education Policy, 16*(2), 149-161. https://doi.org/10.1080/02680930116819

Siegel, M. A., & Wissehr, C. (2011). Preparing for the plunge: Pre-service teachers' assessment literacy. *Journal of Science Teacher Education, 22*(4), 371-391. https://doi.org/10.1007/s10972-011-9231-6

Stiggins, R. (1991). Assessment literacy. *Phi Delta Kappan, 72*(7), 534-539.

Stiggins, R. (1995). Assessment literacy for the 21st century. *Phi Delta Kappan, 77*(3), 238-245.

ISSN 2476-5880

Thomas, L., & Beauchamp, C. (2007). Learning to live well as teachers in a changing world: Insights into developing a professional identity in teacher education. *The Journal of Educational Thought, 41*(3), 229-243.

Tsui, A. B. (2005). Expertise in teaching: Perspectives and issues. In K. Johnson (Ed.), *Expertise in second language learning and teaching* (pp. 167-189). New York: Palgrave Macmillan.

Watmani, R., Asadollahfam, H., & Behin, B. (2020). Demystifying Language Assessment Literacy among High School Teachers of English as a Foreign Language in Iran: Implications for Teacher Education Reforms. *International Journal of Language Testing, 10*(2), 129-144.

Wood, A. (2016). Learning, assessment, and professional identity development in public health training. *Medical teacher, 38*(6), 594-598. https://doi.org/10.3109/0142159x.2015.1073242

Yuan, R. (2019). A critical review on nonnative English teacher identity research: From 2008 to 2017. *Journal of Multilingual and Multicultural Development, 40*(6), 518-537. https://doi.org/10.1080/01434632.2018.1533018

ISSN 2476-5880

Appendix A

Table 1. Exploratory Factor Analysis with Fix Factors: Total Variance Explained

	Initial Eigenvalues			Extraction Sums of Squared Loadings		
Component	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
1	15.21	15.84	15.84	15.21	15.84	15.84
2	6.17	6.42	22.27	6.17	6.42	22.27
3	3.66	3.81	26.09	3.66	3.81	26.09
4	2.89	3.01	29.10	2.89	3.01	29.10
5	2.64	2.75	31.85	2.64	2.75	31.85
6	2.54	2.64	34.50	2.54	2.64	34.50
7	2.49	2.60	37.10	2.49	2.60	37.10
8	2.14	2.23	39.34	2.14	2.23	39.34
9	2.08	2.16	41.51	2.08	2.16	41.51
10	1.91	1.99	43.51	1.91	1.99	43.51
11	1.88	1.96	45.47	1.88	1.96	45.47
12	1.72	1.79	47.27	1.72	1.79	47.27
13	1.66	1.73	49.00	1.66	1.73	49.00
14	1.56	1.63	50.63			
15	1.51	1.57	52.21			
16	1.48	1.54	53.76			
17	1.42	1.48	55.25			
18	1.36	1.42	56.67			
19	1.32	1.38	58.05			
20	1.27	1.32	59.37			
21	1.26	1.31	60.68			
22	1.23	1.28	61.97			
23	1.15	1.20	63.18			
24	1.12	1.17	64.35			
25	1.10	1.15	65.50			
26	1.09	1.13	66.64			
27	1.03	1.08	67.72			
28	1.01	1.05	68.77			
29	.94	.98	69.76			
30	.93	.97	70.73			
31	.91	.95	71.68			
32	.90	.94	72.63			
33	.88	.92	73.55			
34	.85	.89	74.45			
35	.81	.85	75.30			
36	.79	.82	76.12			
37	.77	.80	76.93			
38	.76	.80	77.73			

ISSN 2476-5880

39	.75	.78	78.52
40	.74	.77	79.29
41	.72	.76	80.05
42	.67	.70	80.76
43	.63	.66	81.42
44	.63	.66	82.08
45	.62	.65	82.74
46	.60	.63	83.37
47	.59	.62	83.99
48	.59	.61	84.61
49	.56	.58	85.19
50	.54	.56	85.76
51	.52	.55	86.31
52	.51	.53	86.85
53	.51	.53	87.38
54	.50	.52	87.91
55	.49	.51	88.42
56	.46	.48	88.91
57	.44	.46	89.37
58	.43	.45	89.83
59	.43	.45	90.28
60	.41	.43	90.71
61	.40	.42	91.14
62	.40	.42	91.56
63	.38	.40	91.96
64	.37	.38	92.35
65	.37	.38	92.73
66	.36	.37	93.11
67	.34	.36	93.47
68	.33	.34	93.82
69	.32	.33	94.16
70	.30	.32	94.48
71	.29	.31	94.79
72	.29	.30	95.09
73	.28	.29	95.39
74	.28	.29	95.68
75	.27	.28	95.97
76	.27	.28	96.25
77	.25	.26	96.51
78	.25	.26	96.77
79	.23	.24	97.02
80	.22	.23	97.25
81	.21	.22	97.48

ISSN 2476-5880

82	.20	.21	97.70
83	.20	.21	97.91
84	.20	.21	98.12
85	.19	.20	98.32
86	.18	.19	98.52
87	.17	.18	98.70
88	.17	.18	98.88
89	.16	.16	99.05
90	.15	.16	99.21
91	.14	.15	99.37
92	.13	.14	99.51
93	.13	.13	99.64
94	.12	.12	99.77
95	.11	.11	99.88
96	.10	.11	100.00

Extraction Method: Principal Component Analysis.

ISSN 2476-5880

Appendix B

Table 2. Exploratory Factor Analysis: Rotated Component Matrix

	Component												
	1	2	3	4	5	6	7	8	9	10	11	12	13
VAR00072	.81												
VAR00071	.78												
VAR00073	.73												
VAR00074	.73												
VAR00045	.71												
VAR00046	.66												
VAR00047	.62												
VAR00050	.60												
VAR00082	.60												
VAR00049	.56												
VAR00080	.47												
VAR00048	.41												
VAR00079													
VAR00089		.69											
VAR00090		.67											
VAR00091		.65											
VAR00096		.64											
VAR00088		.62											
VAR00093		.60											
VAR00095		.56											
VAR00092		.56											
VAR00094		.54											
VAR00087		.47											
VAR00026		.47											
VAR00008		.42				.42							
VAR00012		.41											
VAR00025													
VAR00007													
VAR00014			.74										
VAR00015			.61										
VAR00022			.61										
VAR00013			.60										
VAR00024			.58										
VAR00019			.55										
VAR00023			.41										
VAR00037													
VAR00077				.73									
VAR00076				.71									
VAR00075				.68									

ISSN 2476-5880

VAR00078	.44				
VAR00086					
VAR00085					
VAR00033					
VAR00083	.65				
VAR00054	.58				
VAR00053	.55				
VAR00065	.52				
VAR00064	.49				
VAR00063	.42				
VAR00081					
VAR00035					
VAR00084					
VAR00001		.63			
VAR00002		.61			
VAR00005		.55			
VAR00011		.50			
VAR00004					
VAR00021					
VAR00030			.63		
VAR00031			.52		
VAR00029			.50		
VAR00055			.40		
VAR00032					
VAR00016				.54	
VAR00018				.51	
VAR00003				.49	
VAR00010				.45	
VAR00006	.43			.44	
VAR00020				.41	
VAR00017					
VAR00028					
VAR00036					
VAR00068					.74
VAR00067					.67
VAR00066					.48
VAR00070					.48
VAR00060					.43
VAR00062					
VAR00043					.63
VAR00039					.53
VAR00042					.48
VAR00044					.44

ISSN 2476-5880

VAR00061			
VAR00038			
VAR00040			
VAR00051	.56		
VAR00009	.49		
VAR00052	.47		
VAR00056	.43		
VAR00069			
VAR00059		.68	
VAR00058		.63	
VAR00057		.56	
VAR00034			.46
VAR00041			.41
VAR00027			

Extraction Method: PCA.
Rotation Method: Varimax with Kaiser Normalization.
a. Rotation converged in 31 iterations.