

**Construction and Validation of an English Language Teacher Creativity Scale (ELT-CS)**Reza Pishghadam<sup>1</sup>, Purya Baghaei<sup>2</sup>, Shaghayegh Shayesteh<sup>1</sup><sup>1</sup> English Department, Ferdowsi University of Mashhad

PO box 9177948974, Park Square, Ferdowsi University, Mashhad, Iran

<sup>2</sup> English Department, Islamic Azad University, Mashhad Branch, Ostad Yusofi St.  
91871-Mashhad, Iran

**Abstract:** The major purpose of the present study was to construct an English Language Teacher Creativity Scale (ELT-CS). To this end, the questionnaire was designed by picking up the most influential factors of creativity. A total of 24 EFL teachers were rated by their 343 EFL learners on the ELT-CS. The WINSTEPS program was used to perform the Rasch measurement. Due to the fact that the data did not satisfy the criteria for adequate fit to the model, unidimensionality principle was rejected. As a result, consecutive approach was adopted to examine each subscale separately. Individual subscales functioned well and Rasch model held within subscales after removing items 12 and 39. Moreover, since the 5-category rating scale did not operate sufficiently, it was altered to 3-category rating scale. Finally, statistical results were discussed, and implications were provided in the context of English language teaching.

[Reza Pishghadam, Purya Baghaei, Shaghayegh Shayesteh. **Construction and Validation of an English Language Teacher Creativity Scale (ELT-CS)**. Journal of American Science 2012; 8(3):497-508]. (ISSN: 1545-1003). <http://www.americanscience.org>. 66

**Keywords:** Creativity; Rasch measurement; consecutive approach; construct validity

**1. Introduction**

Most recently, a new paradigm in second / foreign language education i.e. life syllabus which simply refers to enhancing the life qualities (e.g. creativity and emotional intelligence) in second / foreign language learning classes, has been pioneered by Pishghadam (2011) and furthered by Pishghadam and Zabihi (2012) in order to draw the researchers' attention to potentials of language learning classes. In fact, language education has shifted its focus from language issues to the educational ones, trying to incorporate the educational and psychological findings in language learning classes. In this regard, it seems that one of the most important issues in psychology which can be dealt in language education is teacher creativity.

Over the past decades, in the world of modern technologies, creativity and innovation have witnessed an overwhelming popularity (Chien, & Hui, 2010; Lau, & Cheung, 2010; Wu, & Albanese, 2010). Likewise, contemporary psychology and pedagogy have found the creationistic approach highly precious; according to which anyone is able to be creative (Karwowski, Gralewski, Lebuda, & Wisniewska, 2007). Given that creative thinking is a key competency for the 21st century, in the first place, experts seek to flourish this fundamental skill by empowering teachers, schools, and educational systems.

This growing need of the society for promoting creative thought has led to what Craft (2005) referred to as 'revolution of creativity in

education'. In accordance, the importance of schooling in cultivating students' creativity is indisputable (Starko, 2010). The classroom has always been an important environment for children to learn how to behave in society. This environment can either encourage or discourage creativity (Eason, Giannangelo, & Franceschini, 2009). However, the environmental factors mainly depend on teachers' perception of creativity education (Chien, & Hui, 2010). As Csikszentmihalyi (1996) indicated that teachers may be important gatekeepers of learners' creative potentials.

A review of the literature demonstrates that, to our best knowledge, no study have been carried out to highlight and ratify the significant role of English as a Foreign Language (EFL) teachers' creativity in their performance in class. Although, there exist several questionnaires for the purpose of measuring this multi-dimensional phenomenon in various fields of studies; none of the tools are merely specified to the realm of language education. To this end, the current study primarily contributes to fill this gap by developing and validating a creativity scale for non-native English language teachers. In fact, the major aim of this questionnaire is to assess how much teachers' activities, strategies, and behaviors in the class cultivate the learners' creative thought.

**2. Theoretical Framework****2.1. Creativity defined**

Unlike numerous concepts in science, there is no single definition of creativity. Nonetheless,

Feldman, Csikszentmihalyi, and Gardner (1994) mutually defined creativity as “the achievement of something remarkable and new, something which transforms and changes a field of endeavor in a significant way” (p.1). Rhodes (1961) estimated nearly 50 definitions of creativity which he qualitatively categorized into four levels known as (4-Ps) model: (a) person, (b) process, (c) press and (d) products. *Person* category indicates information about personality, traits, temperament, and attitudes. This category considers that people with specific characteristics are more creative than others. Rhodes outlined *process* as motivation, learning, and thinking. *Press* hinges on the relationship between human beings and their environment, and *product* is the outcome of a creative endeavor.

As Craft (2001) stated there is a difference between “big C” and “little c” creativity. The former carries a great impact on society whereas the latter concentrates on everyday creativity. Although, its consequence is much less influential, LCC is accepted as the necessary ability of individuals. Furthermore, based on the premises, LCC is considered more relevant to the field of education.

Evidently, the notion of creativity is prominent to teachers due to its direct impact on learning, teaching and more importantly on learners’ future lives. Meanwhile, teachers’ perspective is significantly substantial in the immediate classroom environment. It affects teaching methodologies and educational philosophies. To enable individuals how to treat multiple life ambiguities, teachers need to think beyond the traditional boundaries of launching subject-knowledge. Indeed, this transition necessitates the shift from a traditional subject-teacher to a supportive facilitator of learning (Forrester & Hui, 2007).

Although, working to teach thinking skills and creativity is a true challenge (Hogan, 2006), it has become markedly popular at schools (Burke & Williams, 2008). Ideally, teachers’ role is to empower children to enhance to their highest potential and to lead a more enriching and creative life (Eason et al., 2009). However sadly, nowadays, creativity is implicitly inhibited for the reason that children have to follow plenty of prescribed standards (Hui & Yuen, 2010). Schools teach the learners that knowledge is static and complete; then, instead of producing knowledge, pupils become its great consumers. Nowadays, the primary concern of the teachers and administrators is to have children perform well on standardized tests. As a result, teachers are impelled to engage learners in tedious drills which prepare them for the tests. In this perspective, the significant point is not to overload learners with information but to teach them how to apply it in real life (Laius & Rannikmae, 2003). In other words, students should not be looked at as

empty vessels to be filled; but individuals with creative potential to be cultivated (Eason et al., 2009).

Many attempts have been made in the past to specify various factors contributing to improving creativity. In his study, Kangas (2010) concluded that one way to enhance creativity, imagination and group skill is to amalgamate fact, fiction, and playful learning environment in teaching, learning and studying. Some other experts recognized the role of classroom discussions and cooperative activities beneficial to creative thinking (Beghetto, 2007; Drummond, Mazon, Fernandez, & Wegerif, 2006; Vass, 2007). Furthermore, Eisenberg, Armeli, and Pretz, (1998) and Chien, and Hui (2010) asserted that reward expectancy contributes to creativity development. Besides, Lin (2010) found drama useful in developing certain abilities of creativity such as imagination, risk taking and independent thinking. In another study, Isen, Shalker, Clark and Karp (1978) affirmed that positive emotions nurture creativity by simplifying the access to positive materials in memory.

Fulfilling the demand of our time, educators endeavor to incorporate creativity in school frameworks. Along with the above-mentioned merits, implementing creativity in school curriculums carries some problems as well. Heavy, rigid, and examination oriented curriculums, lack of time, space, confidence, and adequate training were some obstacles before teachers in flourishing creativity at schools (Cheng, 2010). In addition, some students were used to following instructions; thus, they could not deal with free-style creative activities (Cheng, 2010). Furthermore, several teachers complained about the time-consuming nature of creative activities, noise, discipline problems and extra freedom learners had in the class. In line with what is stated, Kampylis, Berki, and Saariluomaa (2009) commented that text books and educational materials do not let learners demonstrate their ultimate sense of creativity.

## 2.2. Creativity measuring tools

The growing interest in creativity has made the researchers develop numerous approaches and tests to measure and evaluate people’s sense of creativity.

Torrance Tests of Creative Thinking (TTCT) (Torrance, 1974) together with the Wallach-Kogan Creativity Tests (WKCT) (Wallach & Kogan, 1965) are probably the two broadly employed measures of creativity. In spite of their popularity, the tests are hard to administer and score; thus, not considered as convenient tools (Lau, & Cheung, 2010). In the same framework, Gough’s Creative Personality Scale (CPS) (Gough, 1979), an adjective checklist, assesses

creative personality. Moreover, the Consensual Assessment Technique (CAT) is regarded as another sort of measurement tool for creativity in which professional experts evaluate creative outcomes. The reason behind its time consuming nature is that in this test participants are asked to draw some pictures or write stories (Amabile, 1982). Creativity Fostering Teacher Index (CFT Index) (Soh, 2000), as a test sharing the most similarities with the present scale, was constructed and validated for teachers to see how much teachers' behaviors contribute to creativity promotion in their students. Paradoxically, the index was a self-rating one and the subjects were merely teachers and not particularly EFL teachers. Despite the noted assessment tools, the researchers' main concern is still accessibility of convenient, reliable and valid measures. Scarcity of such measures would hinder the development of creativity theories and researches (Lau & Cheung, 2010).

On the whole, it seems that the notion of creativity in second or foreign language teaching has not been probed seriously. Available creativity assessments are chiefly designed for the goal of measuring how creative people actually are. It is worth mentioning that the stated measures are substantially designed to be applied to the realm of psychology and not specifically pedagogy. Furthermore, they are commonly composed of lengthy open ended questions and tasks for the subjects to generate their own novel ideas (Silvia, Martin & Nusbaum, 2009). Looking from a pedagogical point of view, it seems that particular attention needs to be paid to this prominent incident. In contrast to what was stated and to compensate for the mentioned deficiencies, the present paper intends to construct an English language teacher creativity scale to determine how much English language teachers are able to *enhance* creativity in their EFL learners, not actually how creative they themselves are.

### 3. Methodology

#### 3.1. Participants

Our community sample consisted of 343 English language learners who rated their 24 English teachers. The learners were both male (N= 100) and female (N=243) with a range of between 16 to 30 years old (Mean= 18.2) in five proficiency levels: elementary, lower intermediate, intermediate, higher intermediate, and advanced. The participants were studying at different private language institutes of Mashhad, Iran. The rationale behind choosing language institutes (informal setting) rather than public schools (formal setting) was the centralized educational policies in Iran's public schools. In this system, decisions are made by the government and teachers are

regarded as mere performers who are impelled to teach the books and materials provided by the people in power. In comparison, educational systems of language institutes are decentralized. Teachers have more freedom in administering their own materials and strategies. Furthermore, in public schools teachers are permanently employed and do not have the fear of losing their jobs; thus, there exists no competition or interest for having a more fruitful class. On the contrary, institute teachers are temporarily employed. This indicates that on the condition they do not absorb more students to the institute; they will be replaced shortly. In all, institute teachers were found to be more eager in conducting different techniques and strategies including creativity, compared to school teachers.

The 24 teachers whom our subjects rated were 10 male and 14 female EFL teachers of the aforementioned institutes aged between 22 and 40 (M= 27.2) with a range of between 2 to 15 (M= 6.7) years of teaching experience. The teachers had all majored in the various branches of English like English teaching, English literature, and English translation at B.A. (N= 9) or M.A. (N=15) level. It is needed to point out that in the educational context of Iran, people educated in diverse branches of English, with an acceptable level of knowledge and proficiency in English language, are allowed to teach English.

#### 3.2. Instrument

An English language teacher creativity scale (ELT-CS) was constructed and then validated through Rasch rating scale model (RSM) (Andrich, 1987) to be used as an instrument for further research in the field of foreign language learning and teaching.

##### 3.2.1. Statistical methodology

The Rasch rating scale model (Andrich, 1978) was adopted as the proper measurement procedure to investigate the data obtained from Likert-type response categories of this scale. In this way, the potential merits of this relatively new theory in comparison with classical test theory (CTT) are notably conspicuous. A major strength of the model is that it rejects the concept of raw scores and provides person and item estimates that are placed on an interval scale. A salient criticism to CTT is that parameters and statistics are test and sample dependent, it means the item difficulties depend on the samples general ability (Ferreira, Almeida, Prieto, 2011). Alternatively, Rasch is item and person free, i.e. it is capable of having estimates for item difficulty and person abilities separately but on a common interval scale. To be specific, it predicts how persons with given underlying ability levels are expected to endorse each item. Further, it mainly focuses on individual items and persons rather than on group statistics (Conrad & Smith, 2004). Rasch model was used in the current paper to substantiate the psychometric properties of the English language

teacher creativity scale. The Rasch rating scale model is formally expressed as follows

$$P(X_{ni}=x) = \frac{\exp \sum_{k=0}^x [\theta_n - (\delta_i + \tau_k)]}{\sum_{k=0}^m \exp \sum_{k=0}^k [\theta_n - (\delta_i + \tau_k)]}$$

$$x=0, 1, \dots, m$$

where  $P(X_{ni}=x)$  is the likelihood that person  $n$  will be observed in rating scale category  $x$  on item  $i$ , which has  $m+1$  rating scale categories,  $\theta_n$  is the persons location on the trait continuum,  $\delta_i$  is the item's difficulty and  $\tau_k$  is the threshold parameter.

The data were analyzed by WINSTEPS Rasch software (Linacre, 2009). Empirical results determine teachers' level of creativity development in their EFL learners.

### 3.2.2. Scale development

Following the taxonomies and guidelines proposed by Rhodes (1961) and Torrance (1974) along with CFT Index (Soh, 2000); plus, delving into plenty of related scientific articles, the researchers came up with the influential factors. They constructed the current scale adopting a straightforward procedure including two steps: 1) Designing the test and 2) Validation.

Applying the instructions provided by the experts in the field of psychology and language teaching, the researchers drafted the questionnaire with 62 items under 7 dimensions in the 5 scale Likert type. Dimensions were titled as originality and elaboration, fluency and flexibility, person (teacher), press (environment) and materials, motivation, independent learning (autonomy), and brainstorming. For each dimension at least 7 items were developed. Joint consultations were also held to revise the items. Afterwards, to assure the content validity of the scale and remove the probable ambiguities of the items 4 EFL teachers and learners were asked to read, think aloud, and suggest improvement for each item. Based upon their comments, we altered the wording of several items to maintain clarity. Finally, near the end of the term 343 EFL learners of several private language institutes in Mashhad, Iran were asked to rate their EFL teachers and fill out the present scale during class hours by prior arrangement with the teachers and administrators.

### 3.3. Procedure

The data was entered into and processed with WINSTEPS software (Linacre, 2009). The validity of the hypothesized factor structure of the ELT-CS questionnaire was examined through Rasch rating scale model (RSM) (Andrich, 1987). In the first place, fit indices together with Principal Component Analysis (PCA) of the residuals were measured to recognize whether the data satisfy the logistic model assumptions or suggest multidimensionality. Afterward, category

functioning was checked to find out if the 5-category rating scale was a suitable choice.

Given that the data did not support unidimensionality principle, consecutive approach was adopted to examine each dimension separately. In essence, consecutive approach is only the unidimensional approach repeated for each dimension independently (Briggs & Wilson, 2003).

Besides, fit estimates, PCA of the residuals and category functioning was similarly investigated for each single dimension. Item and person separation reliabilities were computed as well.

All in all, the scale comprises 62 items. The items are scored according to the Likert type scale of five points ranging from (1) "never" to (5) "always". It must be mentioned that negatively worded items were reverse scored so that a total positively-oriented score can be achieved. That is scoring scheme of items 1, 6, 7, 8, 15, 16, 17, 18, 21, 30, 31, 37, 46, 51, 52, 53, 58 and 59 have been reverse scored.

## 4. Results

This section investigates the evidence for construct validity with respect to dimensionality, item fit, PCA and category functioning. To analyze the data Rasch rating scale model (RSM) (Andrich, 1987), as applied in the software program WINSTEPS (Linacre, 2009), was employed.

### Dimensionality

A primal question in terms of choosing between unidimensional and multidimensional measuring approaches is whether the data consists of one single latent dimension or multiple dimensions. To determine this specific goal, a preliminary analysis was run. The initial analysis of all the 62 items revealed that several items did not fit the model expectations (i.e. unidimensionality). This notion indicates that the items manifesting poor fit would probably measure a construct that is different from the rest of the items or exhibit a further dimension for the concept of creativity.

As Table 1 depicts, items 7, 31, 20, 15, 8, 39 and 6, in descending order of infit mean square (MNSQ) index, broadly misfit the model based on the criteria proposed by Wright and Linacre (1994) for rating scale data (infit MNSQ, 0.6 to 1.4). To specify, the range of mean squares is from 0 to plus infinity with the expected value of 1. Mean squares larger than 1 indicate more variation and smaller than 1 indicate less variation than what the model expects. In accordance, misfit items having infit MNSQ smaller than 0.6 prove overfit, predictability or redundancy; yet, do not contaminate the construct validity of the scale. By contrast, items with MNSQ indices greater than 1.4 deviate from the measurement of the intended construct and degrade the analysis (Linacre, 2005). That is, they

do not measure a single underlying construct uniformly and empirically reject unidimensionality (Baghaei, 2009). Regarding the current table, it should be mentioned that the second column, "Measure", defines item difficulty i.e. the bigger the values, the more difficult the items are. Moreover, "Error", the third column, shows the standard error of the item difficulty measures.

Table 1. Item estimate and fit statistics

Item	Measure	Error	Infit Mean Square
7	-.43	.06	2.12
31	1.23	.05	1.62
20	1.58	.06	1.59
15	-1.10	.08	1.52
8	.53	.04	1.52
39	-.01	.05	1.50
6	-1.13	.08	1.45

Followed by fit MNSQ, Principal component analysis of residuals (PCA), a much more powerful tool, was adopted to confirm multidimensionality. Simply put, PCA is a technique that decomposes the item correlation matrix based on standardized residuals to ratify unidimensionality of the data. Residuals are the discrepancies between the expected probability of correct responses and observed responses. The smaller the discrepancy between the model expectations and observed responses, the better the model and the data fit (Linacre, 2005).

As a matter of fact, in factor analysis of residuals it is not expected to come across a subsidiary dimension; otherwise, the data lack unidimensionality. As Linacre (2006a) maintains, a small eigenvalue of the first PCA component (usually below 2) reveals that residuals are a random noise; on the contrary, a big eigenvalue (usually above 2) proposes a secondary subscale to the intended construct. Based on the results, 1<sup>st</sup> contrast (the largest secondary factor) is estimated 5.3 in eigenvalue units. Furthermore, 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> contrasts are 3.5, 2.4, and 2.2 respectively. Therefore, our teacher creativity scale is by no means unidimensional and clearly embraces four distinct dimensions at least. Nevertheless, they can be regarded as an amalgam of several unidimensional measurement scales into one common scale (Hartig & Hohler, 2009).

On the whole, two decisions could be made upon the results obtained: either delete or revise the irrelevant items to fit a unidimensional measurement model, or consider them as various sub-dimensions of creativity, adopt the consecutive approach and investigate each dimension individually. Eventually, based on the fit statistics and principal component analysis of residuals (PCA), the researchers decided to

run several unidimensional processes analyzing each dimension at a time.

As stated earlier, the fundamental intention of multidimensional approach is to evaluate diverse abilities which contribute to a successful performance in a particular domain (Hartig & Hohler, 2009). The various dimensions of the multidimensional creativity construct can be conceptualized as originality and elaboration, fluency and flexibility, person (teacher), press (environment) and materials, motivation, independent learning (autonomy) and brainstorming.

#### Category Functioning

McCullagh (1980) defined categories as "contiguous intervals on some continuous scale" (p.110). Based on what Bond and Fox (2007) declared the number of categories should be small enough to be distinguishable for the respondent and large enough to cover the whole range of the variable.

As mentioned earlier, our respondents endorsed their level of agreeability based on a 5-point Likert type scale. Relying on Table 2 Rating Scale Statistics, "Category Label" refers to the label of each category and "Observed Count" explains the number of times each category is rated. As Expected, the observed average increases with category values. That is, in this sample higher respondent performance accords with higher categories. "Infit MNSQ" and "Outfit MNSQ" show the average of infit and outfit mean-squares of the responses to each category. Based on Table 2 fit indices do not considerably exceed their acceptable range (0.6 to 1.4).

"Structure Calibration" or Rasch-Andrich thresholds can be interpreted as the points at which the occurrence of either of the two adjacent categories is equally probable (arrows presented in Fig 2) (Linacre, 2005). Given that the first category does not have a preceding category it has no measure and is specified as "None". In threshold perspective, positive values entail that the lower of the two adjacent categories is more likely to be observed; on the other hand, negative values imply that the higher category is more probable to be rated. The numerical ordering of the categories which advance along the continuum is a prerequisite to Rasch measurement. Scrutinizing the table it is found that the step calibration value for category 3 (-0.28) damages the structure's monotonically increasing fashion. Indeed, step calibration disorder might set forth that either the category definition is too narrow, or too many alternatives have been adopted (Linacre, 1999).

Another manifestation of the lack of order is the category probability curves (Fig 1). The curves present the likelihood of a specific category being marked by applicants of diverse ability levels. Theoretically, with regard to what the researchers have

already designed, there should be five visible curves labeled as “5 always”, “4 usually”, “3 sometimes”, “2 rarely” and “1 never”. In principle, on the left side, category 1 (never) is most likely to be noticed for low-ability participants (teachers); whereas, category 5 (always), on the right, is expected to be observed for high-ability participants (teachers). This implies that as the participants’ abilities increase, the probability of rating category 5 enhances similarly. As noted above, the step thresholds prove to be out of sequence (Table 2) and the curves do not demonstrate a natural progression corresponding to a “range of hills” (Fig 1). Due to the narrow interval of the variables, the curves are markedly flat and crowded together. In truth, the distance between the thresholds should not go below 1.4 logits to account for the distinctiveness of the categories and above 5 logits to avoid data loss owing to the absence of appropriate categories (Linacre, 1999). Hence, it can be concluded that our sample respondents were not able to distinguish the difference between the steps and conceptualize the five levels of performance distinctively; in accordance, the 5-categories did not create consistent results.

Table 2. Rating Scale Statistics (12345)

Category Label	Observed Count	Observed %	Infit MNSQ	Outfit MNSQ	Structure Calibration
1	2468	11	1.07	1.19	None
2	2544	12	1.02	1.02	-.22
3	3874	18	.89	.82	-.28
4	5344	25	1.01	.97	.11
5	7263	34	1.01	1.10	.40

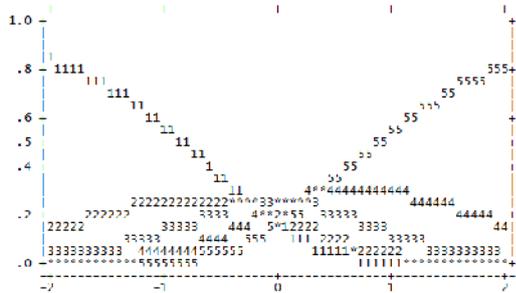


Figure 1. Category probability: Modes\_ Structure measure at intersections (12345)

In order to maintain the category sequence and improve fit to the model, the researchers attempted several configurations including three and four category scales; however, the 12223 pattern produced the best indices. Consequently, they decided to integrate the adjacent categories 2, 3 and 4 into 2 and shift the 12345 category pattern (5-point scale) to 12223 (3-point scale) (Table 3, Fig 2).

Table 3 reports that the overall rating scale statistics operate satisfactorily after collapsing categories. The observed average and structure

calibration empirically increase along with category counts. Furthermore, the threshold parameters together with mean square fit indices fall into their acceptable range of 1.4 to 5 and 0.6 to 1.4, respectively.

Table 3. Rating Scale Statistics (123)

Category Label	Obsvd Avreg	Sample Expect	Infit MNSQ	Outfit MNSQ	Structure Calibration
1	-.61	-.70	1.07	1.11	None
2	.54	.58	.92	.89	-1.58
3	1.64	1.61	.99	1.03	1.58

Corresponding to Table 3, Figure 2 illustrates the category curves in expected succession of "hills". That is, the step calibrations become more positive as the category values advance. Despite the fact that 5-category rating scale structure was intended primarily, in reality merely three categories functioned properly.

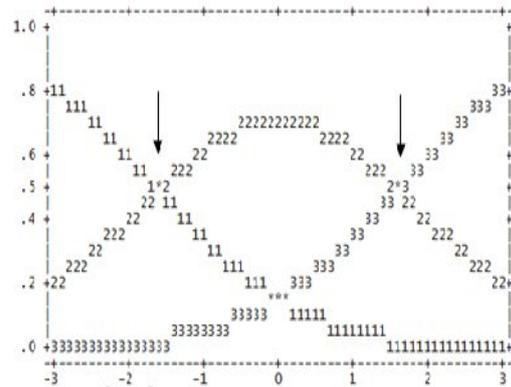


Figure 2. Category probability: Modes\_ Structure measure at intersections (123)

As multidimensionality was formerly demonstrated, in the following we are going to adopt a consecutive approach and substantiate the validity of each single dimension through investigating fit indices for the items, reliability and separation statistics along with principal component analysis of residuals and category diagnostics.

*Consecutive Approach*

Table 4 summarizes the statistical findings (including reliabilities, separation, error of measurement, standard deviation and item measurement for items and persons) related to each single subscale.

Analyzing the items within their pertinent scales, it was concluded that all fulfilled the Rasch model expectation excluding items 12 and 39 (Table 5). In order to improve model fit, misfitting items were removed from the measurement owing to the fact that they might measure a construct that was different from the other items of the intended dimension. As formerly

mentioned, the unidimensional analysis of the 62 items resulted in the misfit of 7 items and an unexplained variance in contrast 1 of 5. Yet, when the items are analyzed within their relevant subscales, item fit improves in a way that merely two items misfit and all unexplained variances are smaller than 2 or 2. This notion supports the hypothesis that creativity is best modeled as a multidimensional construct.

Concerning PCA of the residuals, unexplained variances in the first contrasts for all the seven scales were 2 or lower than 2 which exhibit that the subscale were unidimensional. Table 6 contains the unexplained variance of the first contrasts for the seven subscales.

**Table 4.** Summaries of measured items and persons

Dimension		Reliability	Separation	RMSE	SD	Measures	
Originality & Elaboration	Item	.96	5.09	.12	.63	Min	-1.40
						Max	.82
Person	Person	.74	1.68	.16	.80	Min	.67
						Max	3.71
Fluency & Flexibility	Item	.94	4.00	.11	.43	Min	-.74
						Max	.73
Person	Person	.81	2.00	.07	.63	Min	-.03
						Max	2.90
Person	Item	.99	9.55	.11	1.05	Min	-1.69
						Max	1.36
Person	Person	.77	1.80	1.7	.46	Min	.37
						Max	2.13
Press & Materials	Item	.99	13.44	.10	1.35	Min	-2.03
						Max	1.69
Person	Person	.76	1.77	.05	.46	Min	-1.14
						Max	.92
Motivation	Item	.98	7.60	.11	.86	Min	-1.31
						Max	1.84
Person	Person	.70	1.52	.11	.64	Min	-.65
						Max	1.96
Autonomy	Item	.99	9.62	.11	1.06	Min	-1.57
						Max	1.99
Person	Person	.74	1.68	.08	.09	Min	-1.02
						Max	1.42
Brain storming	Item	.99	11.66	.11	1.34	Min	-2.13
						Max	1.96
Person	Person	.77	1.80	.08	.63	Min	-.25
						Max	2.15

**Table 5.** Item estimate and fit statistics

Dimension	Item	Measure	Error	Infit MNSQ
Originality & Elaboration	12	.82	.12	<b>1.42</b>
	13	-1.40	.13	.84
	26	-.19	.12	.93
	27	-.14	.12	.78
	48	.11	.12	.81
	50	.02	.12	.85
	51	.77	.12	1.04
	59	.01	.12	1.27
	Fluency & Flexibility	2	.73	.10
22		-.41	.11	.84
23		.44	.10	1.23
24		-.30	.11	.81
25		.45	.10	1.12
28		.28	.10	1.07
29		-.12	.11	.84
49		.26	.10	.82
52		-.74	.11	1.18
53		-.21	.11	1.20
56		.20	.11	.76
63	-.58	.11	1.17	
Person	5	.82	.10	.67

Originality & Elaboration	6	-1.69	.13	1.05
	21	1.22	.10	1.6
	30	-.26	.11	1.10
	32	-1.52	.12	.89
	39	.13	.10	<b>1.57</b>
	45	.20	.10	.81
	46	1.36	.10	.81
	57	-1.07	.11	.96
	58	.81	.10	1.10
Press & Materials	7	-2.03	.10	.42
	8	-.14	.09	.82
	14	.02	.09	.98
	20	1.69	.11	1.01
	31	1.01	.10	1.05
	36	1.25	.10	.87
	61	-1.79	.10	.93
Motivation	9	.09	.11	1.12
	10	1.84	.11	1.24
	19	-1.31	.12	1.03
	33	-.49	.11	.71
	34	.41	.11	.98
	38	-.88	.11	1.00
	42	.68	.11	1.22

	44	.57	.11	.85
	54	-.54	.11	.83
	60	-.37	.11	.96
Autonomy	11	-.60	.11	.87
	17	-.24	.11	.65
	18	-.93	.11	1.02
	35	-.11	.11	.84
	37	-1.57	.11	1.33
	41	1.13	.11	1.24
	43	1.99	.11	1.21
Brain storming	55	.34	.11	.83
	1	-1.45	.12	.89
	3	-.17	.11	.99
	15	.213	.14	1.08
	16	1.96	.11	1.34
	40	.28	.11	.82
	47	.09	.11	.90
	62	1.42	.11	.96

Table 6. Table of standardized residual variance

Dimension	Unexplained variance - 1st contrast In eigenvalue units
Dimension 1	1.6
Dimension 2	1.8
Dimension 3	2.0
Dimension 4	2.0
Dimension 5	1.5
Dimension 6	1.9
Dimension 7	1.7

Based on Table 7 the category thresholds were all ordered and their fits were close to the perfect value of 1. Therefore, the 3-category scale

Table 7. Rating scale statistics

Dimension	Category Label	Infit MNSQ	Threshold
Originality & Elaboration	1	1.10	None
	2	.93	-2.36
	3	1.00	2.36
Fluency & Flexibility	1	1.07	None
	2	.94	-1.70
	3	.99	1.70
Person	1	1.14	None
	2	.92	-1.74
	3	.98	1.74
Press & Materials	1	1.04	None
	2	.93	-1.24
	3	1.02	1.24
Motivation	1	1.00	None
	2	.93	-1.98
	3	1.3	1.98
Autonomy	1	1.05	None
	2	.90	-1.91
	3	1.01	1.91
Brainstorming	1	1.19	None
	2	.92	-1.89
	3	.92	1.89

## 5. Discussion

To investigate the extent to which EFL teachers foster creativity in their EFL learners 62-item scale (ELT-CS) was designed. Rasch model was applied to substantiate the construct validity of the scale in the context of EFL teachers and learners.

The results of the initial Rasch measurement indicated violation of the unidimensionality principle. Thus, consecutive approach was adopted to examine each dimension separately under unidimensional Rasch rating model. In all, it was found that except for items 12 and 39, the remaining items, functioned well under seven subscales; each subscale measures a single latent dimension. To achieve satisfactory fit it was essential to remove Item 12 from Dimension 1, "Originality and Elaboration", and Item 39 from Dimension 3, "Person". Verifying the content of the two stated items it was perceived that the statistical findings were not in line with the previously proved notions of some experts in earlier studies.

Item 12 is the teacher's preference of using open-ended questions to multiple-choice ones. Relying on the infit MNSQ of above 1.4, the item is distinguished to be irrelevant to the intended construct. The results obtained were not incompatible with Bredekamp and Copple's (1997) idea that a creative classroom should devote more time to open-ended questioning. The first probable reason may be that the question was slightly vague and the participants were not able to grasp whether the exact purpose is the tests or the classroom questions that the teacher generally asks. The second likely reason may refer to the fact that, institute teachers are not totally free in how to investigate their learners' level of knowledge. For the ease of administration and scoring, English language institutes mainly design the midterm and final exams in multiple-choice forms irrespective of the teacher, the textbooks and the course goals. According to such educational policies, the relevance of the item to the concept of teacher's role in fostering creativity and producing original ideas is ignored.

Further, item 39 is concerned with listening to a conversation for the first time while the books are closed. In particular, the basic aim is to improve learners' sense of imagination. Based on Vygotsky's theory, there is a powerful link between imagination and creativity; as a matter of fact, imagination is the essence of creativity (Eckhoff & Urbach, 2008; Lindqvist, 2003). Yet, our empirical evidence revealed that this item does not connect with other items of the current scale. That is, the item estimates another concept rather than teacher's role in creativity. Psychologically, it is possible to say that anyone who is imaginative is not creative. In general sense, there are times that creativity departs from imagination. If people merely make images in their minds and do not

generate an output or give expression to that, then they are perhaps more imaginative than creative.

Besides, it was discovered that mistakenly items 4 and 56 were repeated. In reference, item 4 was randomly omitted from the scale. Ultimately, after removal of the 3 items, the total number of items comprising the scale was equal to 60.

To boot, according to psychometric evidence obtained from the functioning of response categories, the 5-category structure (12345), which was chosen primarily, did not function effectively for the ELT-CS. It is interesting to know that the statistical results corroborate participants' similar feeling of dissatisfaction. During the pilot study, respondents constantly complained about the categories' resemblance and had a hard time selecting from among the 5 choices. Alternatively, after inspecting multiple configurations, 3-category structure (12223) proved to be a better option. As a result, in order to maintain distinctiveness, alternation was required in the original rating scale. Simply put, choices 2 (usually), 3 (sometimes) and 4 (rarely) combined into one and the rating pattern shifted to: 1 (always), 2 (sometimes) and 3 (never).

The analysis manifested the hierarchy of creativity fostering behaviors in EFL teachers. The most difficult items were found to be 43 and 10. These items were "*asks successful learners to talk about their learning strategies in class*", and "*more than one topic is offered to us to choose for the writing*", respectively. In truth, the noted items require a higher degree of creativity fostering behavior of the teachers. Basically, considering fairness and ease of scoring only one topic is given to the learners to write.

On the other hand, items 15 and 7 were the easiest ones. These negatively scored items were "*she mocks learners' seemingly irrelevant ideas*" and "*we use supplementary books along with our main text book*". The items imply that a lower degree of creativity fostering behavior is needed for the teachers to perform the stated activities in the class. According to our statistical evidence majority of teachers were rated high on these two items. In general sense, teachers rarely ridicule their learners' ideas or make use of extra books. Logically, depending on English language institutes regulations and policies, teachers are recommended not to administer supplementary books and cover the main book merely.

The value of this questionnaire lies in constructing an ELT-CS which can be of great help to researchers interested in studying creativity and institute managers in recruiting eligible teachers who are able to identify and cultivate learners' creative potential. Similarly, the test can be applied to the current teachers of English language teaching centers. Based on the empirical statistics and in accordance

with Chien and Hui's (2010) conception, appropriate training courses would be planned in practice to elevate teachers' creativity knowledge and education. Over and above, the content of the items can exclusively serve miscellaneous ways of improving English teachers creativity fostering behaviors. By using creativity developing techniques and realizing their role in bringing change in learning and teaching context, the teachers can make progress in achieving their ultimate capabilities.

After all, further research in larger and appropriately targeted samples is recommended to support and improve the current instrument. Also, several studies could be conducted using this English language teacher creativity scale to find its objective association with various pedagogical and psychological variables. Another possible research could focus on ways to improve teachers' tendency toward implementing more creative methods and tasks in the classroom.

Manifestly, readers must keep in mind that a study such as the present one has its own restrictions. Verifying predictive validity of the scale as a necessary aspect was postponed to the future studies since it would go beyond the scope of this research. At last, the participants involved in this study were selected from a number of English language institutes which were not representative of the big population of English language teachers and learners of neither Iran nor Mashhad. In reference, universal generalization of the findings is not recommended; yet, the implication of the data might be useful for similar contexts and samples.

## References

1. Amabile, TM. Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology* 1982; 43(5): 997-1013.
2. Andrich, D. A rating formulation for ordered response categories. *Psychometrika* 1978; 43(4): 561-573.
3. Baghaei, P. *Understanding the Rasch model*. Mashhad Islamic Azad University Press, Mashhad, 2009.
4. Beghetto, RA. Does creativity have a place in classroom discussions? Prospective teachers' response preferences. *Thinking Skills and Creativity* 2007; 2: 1-9.
5. Bond TG, Fox CM. *Applying the Rasch model: Fundamental measurement in the human sciences*. Lawrence Erlbaum, London, 2007.
6. Bredekamp S., Copple C. (Eds.) *Developmentally appropriate practice in early*

- childhood programs. NAEYC, Washington, 1997.
7. Briggs DC, Wilson M. An introduction to multidimensional measurement using Rasch models. *Journal of Applied Measurement* 2003; 4(1): 87-100.
  8. Burke LA, Williams JM. Developing young thinkers: An intervention aimed to enhance children's thinking skills. *Thinking Skills and Creativity* 2008; 3: 104-124.
  9. Cheng VMY. Tensions and dilemmas of teachers in creativity reform in a Chinese context. *Thinking Skills and Creativity* 2010; 5: 120-137.
  10. Chien C, & Hui AN. Creativity in early childhood education: Teachers' perceptions in three Chinese societies. *Thinking Skills and Creativity* 2010; 5: 49-60.
  11. Conrad KJ, Smith EV. International conference on objective measurement: Applications of Rasch analysis in health care. *Medical Care* 2004; 42(1): 1-6.
  12. Craft A. Little c creativity. In A Craft, B Jeffrey, M Liebling (Eds.), *Creativity in education*. Continuum, London, 2001.
  13. Craft A. *Creativity in schools: tensions and dilemmas*. Routledge, London, 2005.
  14. Csikzentmihalyi M. *Creativity: Flow and the psychology of discovery and invention*. Harper Collins, New York, 1996.
  15. Drummond SR, Mazon N, Fernandez M, Wegerif R. Explicit reasoning, creativity and co-construction in primary school children's collaborative activities. *Thinking Skills and Creativity* 2006; 1(2): 84-94.
  16. Eason R, Giannangelo DM, Franceschini LA. A look at creativity in public and private schools. *Thinking Skills and Creativity* 2009; 4: 130-137.
  17. Eckhoff A, Urbach J. Understanding imaginative thinking during childhood: Sociocultural conceptions of creativity and imaginative thought. *Early Childhood Education Journal* 2008; 36(2): 179-185.
  18. Eisenberg R, Armeli S, Pretz J. Can promise of reward increase creativity? *Personality Processes and Individual Differences* 1998; 74(3): 704-714.
  19. Feldman DH, Csikszentmihalyi M., Gardner H. A framework for the study of creativity. In *Changing the world: A framework for the study of creativity*. Westport, CT, Praeger, 1994.
  20. Forrester V, & Hui, A. Creativity in the Honk Kong classroom: What is the contextual practice? *Thinking Skills and Creativity* 2007; 2: 30-38.
  21. Frreira AI, Almeida LS, Prieto G. The role of processes and contents in human memory: an item response theory approach. *Journal of Cognitive Psychology* 2011; 23(7): 873-885.
  22. Gough HG. A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology* 1979; 37(8): 1398-1405.
  23. Hartig J, Hohler J. Multidimensional IRT models for the assessment of competencies. *Studies in Educational Evaluation* 2009; 35: 57-63.
  24. Hogan MJ. Cultivating thought-full and creative thinkers: A comment on quality merging Costa with Claxton et al. *Thinking Skills and Creativity* 2006; 1: 152-154.
  25. Hui ANN, Yuen TCM. The blossoming of creativity in education in Asia: Changing views and challenging practices. *Thinking Skills and Creativity* 2010; 5: 155-158.
  26. Isen AM, Shalke TE, Clark M, Karp L. Affect, accessibility of material in memory, and behavior: A cognitive loop. *Journal of Personality and Social Psychology* 1978; 36(1): 1-12.
  27. Kamylyis P, Berki E, Saariluomaa P. In-service and prospective teachers' conceptions of creativity. *Thinking Skills and Creativity* 2009; 4: 15-29.
  28. Kangas M. Creative and playful learning: Learning through game co-creation and games in a playful learning environment. *Thinking Skills and Creativity* 2010; 5: 1-15.
  29. Karwowski M, Gralewski J, Lebeda I, Wisniewska, E. Creative teaching of creativity teachers: Polish perspective. *Thinking Skills and Creativity* 2007; 2: 57-61.
  30. Laius A, Rannikmae M. Influence of STL teaching on students' creativity. *Science Education International* 2003; 14(4): 21-28.
  31. Lau S, Cheung PC. Creativity assessment: Comparability of the electronic and paper-and-pencil versions of the Wallach-Kogan Creativity Tests. *Thinking Skills and Creativity* 2010; 5: 101-107.
  32. Lin YS. Drama and possibility thinking – Taiwanese pupils' perspectives regarding creative pedagogy in drama. *Thinking Skills and Creativity* 2010; 5: 108-119.
  33. Linacre JM. Category disordering vs. step (threshold) disordering. *Rash Measurement Transactions* 1999; 13(1): 675. <http://www.rasch.org/rmt/rmt131a.htm>

34. Linacre JM. WINSTEPS Rasch programme. Version 3.63.0, 2009.
35. Linacre JM. A user's guide to WINSTEPS/MINISTEP: Rasch-model computer programs. Winsteps.com, Chicago, IL, 2006.
36. Lindqvist G. Vygotsky's theory of creativity. *Creativity Research Journal* 2003: 15(2-3): 245-251.
37. McCullagh P. Regression models for ordinal data. *J. Royal Statist. Soc B* 1980: 42(2): 109-142.
38. Pishghadam R. Introducing Applied ELT as a new approach in second/foreign language studies. *Iranian EFL Journal* 2011: 7 (2): 8-14.
39. Pishghadam R., Zabihi R. Life syllabus: A new research agenda in English language teaching. *Perspectives* 2012.
40. Rhodes M. An analysis of creativity 1961. *Phi Delta Kappan*: 42(7): 305-310.
41. Soh KC. Indexing creativity fostering teacher behavior: A preliminary validation study. *Journal of Creative Behavior* 2000: 34(2): 118-134.
42. Starko AJ. *Creativity in the classroom: Schools of curious delight* (4th ed.). Routledge, New York, 2010
43. Torrance EP. *Torrance tests of creative thinking*. Personnel Press, Lexington, MA, 1974.
44. Vass E. Exploring processes of collaborative creativity—The role of emotions in children's joint creative writing. *Thinking Skills and Creativity* 2007: 2: 107-117.
45. Wallach MA, Kogan N. *Modes of thinking in young children: A study of the creativity intelligence distinction*. Holt, Rinehart & Winston, New York, 1965.
46. Wright BD, Linacre JM. Reasonable mean-square fit values. *Rasch Measure Transactions* 1994: 8(3): 370.
47. Wu J, Albanese D. Commentary Asian creativity, chapter one: creativity across three Chinese societies. *Thinking Skills and Creativity* 2010: 5: 150-154.

## Appendix

### The ELT-CS: 62 items

No	Statement	Always	Usually	Sometimes	Rarely	Never
1	Interrupts the learners while expressing their ideas.	1	2	3	4	5
2	Asks us synonyms and antonyms.	5	4	3	2	1
3	We are required to <i>guess</i> the meaning of the new words in the first place.	5	4	3	2	1
5	Administers various teaching methods.	5	4	3	2	1
6	Gets tired of our numerous questioning.	1	2	3	4	5
7	We use supplementary books along with our main text book in the class.	1	2	3	4	5
8	Assigns several rules for the class to obey.	1	2	3	4	5
9	Appreciates our both right and wrong responses.	5	4	3	2	1
10	More than a single topic is offered to us to choose for each writing task.	5	4	3	2	1
11	Teaches us how to learn more effectively.	5	4	3	2	1
12	Teacher's questions are mainly open-ended rather than multiple-choice.	5	4	3	2	1
13	Cares a lot for class discussions.	5	4	3	2	1
14	Makes use of flash cards and videos in her teaching.	5	4	3	2	1
15	Mocks learners' seemingly irrelevant ideas.	1	2	3	4	5
16	Comments on the truthfulness of our responses on the spot.	1	2	3	4	5
17	Talks more than the learners in the class.	1	2	3	4	5
18	Answers different questions immediately without getting help from us.	1	2	3	4	5
19	Values our learning more than our grades.	5	4	3	2	1
20	We are allowed to walk and move in the class.	5	4	3	2	1
21	Reminds us dos and don'ts.	1	2	3	4	5
22	After teaching new grammatical points, helps us to make similar examples.	5	4	3	2	1
23	Before starting a new conversation or reading we should guess the theme from the provided pictures.	5	4	3	2	1
24	Accepts learners' ideas that contradict his/hers.	5	4	3	2	1
25	Asks us to talk about our favorite topics for a couple of minutes.	5	4	3	2	1
26	Encourages our novel, original ideas.	5	4	3	2	1
27	Helps us to be clear in discussions.	5	4	3	2	1
28	Learners who comment more are encouraged more.	5	4	3	2	1
29	We are required to put the learned materials into use.	5	4	3	2	1
30	The learners who do not observe the class rules are punished.	1	2	3	4	5
31	Insists on carefully covering the whole book.	1	2	3	4	5
32	Necessitates learning the basic materials accurately.	5	4	3	2	1
33	Applies our favorite topics in the class as far as possible.	5	4	3	2	1
34	Chooses writing topics that are closely related to everyday life.	5	4	3	2	1

35	We are expected to check our work before s/he does.	5	4	3	2	1
36	We play different games in the class.	5	4	3	2	1
37	Writes the meaning of the new words on the board without asking learners' interpretations.	1	2	3	4	5
38	Some of the exercises are done in groups.	5	4	3	2	1
39	Asks us to listen to a conversation for the first time while our books are closed.	5	4	3	2	1
40	Before starting a new topic s/he reviews our background knowledge.	5	4	3	2	1
41	We read our writings in the class for our classmates.	5	4	3	2	1
42	Competitions are chiefly cooperative rather than individual.	5	4	3	2	1
43	Asks successful learners to talk about their learning strategies.	5	4	3	2	1
44	Mentions the goal of each exam or exercise.	5	4	3	2	1
45	Takes our opinions serious and follows them up.	5	4	3	2	1
46	His/her behavior in class is predictable.	1	2	3	4	5
47	Points to the title of each section and appreciates us to guess the subject.	5	4	3	2	1
48	Encourages learners' original and novel interpretations.	5	4	3	2	1
49	Asks the similarities and differences of the pictures, sentences and texts.	5	4	3	2	1
50	Asks questions to make us think deeper.	5	4	3	2	1
51	Asks us repetitive questions.	1	2	3	4	5
52	Learners are allowed to give <i>only</i> one response to teacher's questions.	1	2	3	4	5
53	According to him/her, questions constantly have one correct answer.	1	2	3	4	5
54	We are allowed to talk about our experiences in the class.	5	4	3	2	1
55	Some questions are left unanswered for us to explore.	5	4	3	2	1
56	Examples of grammatical points are related to everyday life.	5	4	3	2	1
57	Listens carefully to our questions and answers.	5	4	3	2	1
58	Takes exams regularly.	1	2	3	4	5
59	Interprets the text irrespective of our opinions and interpretations.	1	2	3	4	5
60	Chooses learners' favorite topics for class discussions.	5	4	3	2	1
61	Keeps the atmosphere of the class happy.	5	4	3	2	1
62	To facilitate the process of writing, teacher reviews our background knowledge and writes them on the board in categories.	5	4	3	2	1
63	After covering each conversation, s/he expects us to make a change or create a new conversation based on our own situation.	5	4	3	2	1

2/15/12