

*Journal of Teaching Language Skills (JTLS)*  
36(1), Spring 2017, pp. 111-144- ISSN: 2008-8191  
DOI: 10.22099/jtls.2017.4045

## **An Assessment Scheme for ELT Performance :An Iranian Case of Farhangian University**

Gholamreza Kiany                      Mohammad Nabi Karimi  
Associate Professor                      Associate Professor  
Tarbiat Modares University              Kharazmi University  
rezakiany@yahoo.com                      karimi\_mn@yahoo.com

Monireh Norouzi \*  
PhD Student  
Tarbiat Modares University  
norouzimonireh@gmail.com

### **Abstract**

Accountability concerns in language education call for the development of more valid and authentic measures of assessment. In light of these concerns, performance assessment has received increasing interest in the context of teacher education programs and teacher licensing over the last decade. In Iran, a recent policy adopted by Farhangian University aims at assessing the professional competencies of its ELT graduates by requiring them to go through a performance assessment as part of the licensing requirements. Mounting concerns regarding the validity of traditional tests used for teacher certification (Mitchell, Robinson, Plake, & Knowles, 2001) have motivated Farhangian University to develop its own performance assessment. Therefore, the present study explored the components of the performance assessment through detailed analysis of the Curriculum Document of the English Major, review of literature, and investigation of the stakeholders' perspectives. To this end, in this exploratory study, convenience, purposive, and cluster sampling procedures were used for the selection of the teacher educators, student-teachers, and mentor teachers. Then, in-depth interviews were conducted with the stakeholders. Finally, based on the content analysis of the above-mentioned sources which resulted in a strong agreement, a performance assessment scheme with seventeen items was developed. However, results of the factor analysis yielded a thirteen-factor performance assessment scheme to be used as the criterion for assessing the professional competencies of student-teachers.

*Keywords:* assessment, performance assessment, accountability

---

Received: 24/10/2016      Accepted: 29/04/2017

\* Corresponding author

Living in an era of accountability requires educational systems to ensure quality education. Since learners' academic achievement is the best indicator of quality education (Hodge & Morgan 2014; Sanders, Wright, & Horn, 1997), a number of studies have been conducted to analyze the factors that have the largest share in learners' academic success. What emerges from these studies is that teachers are an integral part of highly-qualified education (Gasper & Vieira 2013; Hodge & Morgan 2014; Sanders, Wright, & Horn, 1997).

The recognition of the significant influence of teachers on learners' academic achievement and on the success of any educational system comes with the need to find the most reliable way of evaluating teacher performance and effectiveness (Medley, 1982). Teacher evaluation is a complex concept that involves a number of measures affecting teachers and teaching. It refers to "the systematic assessment of a teacher's performance and/or qualifications in relation to the teacher's defined professional role and the school district mission" (Shinkfield & Stufflebeam, 1995, p. 86). The need for teacher evaluation is also endorsed by Beare (1989) who maintains that "teacher assessment will always be needed; any enterprise or activity needs assessment, review, and constant searching for better practices; any professional operator needs that kind of informed feedback too" (p. 10).

The need for teacher evaluation as a tool contributing to professional development, learners' academic achievement, and quality assurance has revealed the potential benefits of high-stakes assessments, and in particular, performance-based assessments, for teacher learning, teaching quality, and student achievement (Danielson & Marquez, 1998; Delandshere & Arens, 2003; Haertel, 1991). Since performance assessment provides contextualized evidence of student learning, they are more powerful tools to predict teachers' roles in learners' achievement gains (Darling-Hammond, 2010). Besides being useful for students, the literature sets out several convincing reasons that performance assessments advantage teachers, too. Sandholtz (2012), for instance, maintains that performance assessment indicates how teachers use their

knowledge and skills in their teaching practices, reflect on their practices, and consequently make use of appropriate instructional strategies to achieve effective teaching. They result in teachers' use of more teaching strategies, support of student learning, and modification of their practices based on assessment results (Darling-Hammond, 2010). A similar argument is presented by Chung (2007), who argues that performance assessment enables teacher candidates to modify their instruction based on assessment results from the perspective of student learning. Pecheone and Chung (2006) hold that "performance assessments that include evidence from actual teaching practice have the potential to provide more direct evaluation of teaching ability" (p. 23). Darling-Hammond (2010) has also acknowledged the role of performance assessment in increasing teachers' subject matter knowledge, improving classroom management, and designing instruction. Particularly important, performance assessments help novice teachers through "changing their understanding of teaching and their practice" (Darling-Hammond, 2010, p. 14). In a similar vein, Chung (2007) have found that performance assessments exert a positive influence on the professional learning of teachers.

Apart from having an impact on individual teacher candidates, performance assessments for licensing do benefit teacher education programs which is equally important (Selvester, Summers, & Williams, 2006). Teacher preparation programs can improve curriculum and program design through using data obtained from performance assessments. Following a questionnaire-based survey, Selvester et al. (2006), for instance, realized candidates' needs to be supported through mentoring and direction during performance assessments. Consequently, attempts were made by the faculty members to improve the articulation of their program courses which finally led to improvements in the teacher education program. Specifically, the data can also be examined "for program accreditation to provide a basis for deciding which programs should be encouraged, improved, or closed if they cannot improve enough to enable most of their candidates to demonstrate that they can teach" (Darling-Hammond, 2010, p. 17). Adopting a different

perspective, the scoring process contributes to the improvement of teacher preparation programs, as well. As stated by Darling-Hammond (2010), "the act of scoring is itself also educative" (p. 18). Since performance assessments are assessed by faculty members and supervisors, in order to assess candidates' performances based on standardized rubrics, they are required to participate in training courses, which in turn, improve teacher preparation programs (Darling-Hammond, 2010).

Although the importance of teacher evaluation is acknowledged by a number of researchers, the current teacher evaluation methods have come under fire for being ineffective (McGreal, 1988; Peterson, 2000; Prybylo, 1998). As suggested by Danielson and McGreal (2000), the evaluative criteria should reflect the latest findings of educational research. However, many evaluation methods are reported to suffer from serious deficiencies such as utilizing outdated, limited evaluative criteria; indicating few shared values and assumptions about good teaching; and lacking precision in evaluating performance (Danielson & McGreal, 2000). They are also criticized for their disregard of context, process, and reflection (Bunch, Aguirre, & Téllez, 2009). More specifically, current evaluation methods used in the Iranian educational systems do not lead to quality education, learners' academic achievement and professional development or in other words, they fail to accomplish the purposes for which they are used (Navidinia, Kiani, Akbari, & GhafarSamar, 2013). In this respect, this study is an attempt to address some of the criticisms leveled at previous teacher evaluation methods used in Iran's context by proposing a performance assessment scheme.

### **Conceptual Framework**

The profound effect of teachers on quality education has led to repeated attempts to explore factors contributing to teacher quality (Sanders, Wright, & Horn, 1997, p. 57). This line of inquiry has resulted in various definitions of competency. Shulman (1987), for instance, refers to competency as a combination of knowledge, skills, and attitudes

appropriate to the context. He further argues that teachers must be familiar with the core knowledge of teaching which refers to the knowledge of teaching elements. Interestingly, a close examination of the Fundamental Educational Reform Document indicates that teacher education within Farhangian University is competency oriented within which competency is defined as a set of qualities, abilities, and skills which is to be acquired by teachers in order to bring about changes in their understanding and performance. Based on this document, teacher professional competency as a "core competency" consists of key competencies and basic competencies. The major dimensions of such an interpretation of teachers' professional competencies which play fundamental roles in the development of the curriculum of Farhangian University entails content knowledge (CK), pedagogical knowledge (PK), pedagogical content knowledge (PCK), and general knowledge (GK).

Concomitant to investigating and improving teaching quality is the use of comprehensive assessments. Developing teacher evaluation systems depends on a solid understanding of the task of teaching (Taut & Sun, 2014). In other words, it is impossible to think of teacher evaluation without directly investigating the nature of teaching. Therefore, there is a widespread agreement among researchers (e.g. Berliner 2005; Darling-Hammond & Snyder 2000; Haertel 1991; Shulman 1986) that there is a need to design assessment tools to assess different competencies including pedagogical content knowledge and that it is impossible to measure teachers' knowledge completely and appropriately without assessing pedagogical content knowledge. Based on Shulman's (1986) description of pedagogical content knowledge as "the most regularly taught topics in one's subject area, the most useful forms of representation of those ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations" (p. 9), Koirala, Davis, and Johnson (2008) also conclude that pedagogical content knowledge plays a significant role in student achievement gains, thus any assessment

aimed at measuring teachers' knowledge, apart from content knowledge, must assess pedagogical content knowledge as well.

The pivotal role of teachers in quality education and the need for evaluating teachers' competencies, and more specifically, teachers' pedagogical content knowledge has paved the way for the increased use of performance assessments in teacher education and teacher certification programs (Darling-Hammond, 2010). Therefore, the body of research on performance assessment, referring to "a measure of the complex pedagogical skills required for candidates to successfully teach and cause their students to learn" (Torgerson, Macy, Beare, & Tanner, 2009, p. 80), has burgeoned over the last decades. Motivated by the same concerns, Farhangian University has adopted a new policy, *ASLAH* (Evaluation of Professional Competencies), which requires student-teachers to go through a performance assessment prior to certification.

**Objectives.** The importance of teacher evaluation in Iran's educational system, in addition to what was mentioned above, arises from the new project proposed by Farhangian University. Required by this new project, *Assessment of the Professional Competencies*, all student-teachers must go through this comprehensive assessment system prior to obtaining teaching licensure. This project aims to increase the quality of education by improving the professional development of its ELT graduates. For this purpose, performance assessment, written assessment, portfolio, and GPA are used as the criteria for evaluating these competencies. In light of this background, there is an urgent need to conduct studies to explore different components of the new assessment system. However, no research is conducted so far in this regard probably due to its recent promulgation. Therefore, well designed pilot research that can provide insights into the development and implementation of each part of the project would contribute to the efficiency of the whole project.

The present study aims at providing initial insights into the development of performance assessment by exploring its components through detailed analysis of the Curriculum Document of the English

major, review of literature, and interviews with stakeholders, and consequently developing a performance assessment scheme. Specifically, the study aims to address the following research question:

*What are the major factors of a teacher performance scheme to be used as a benchmark for assessing Farhangian University's ELT student-teachers' competencies upon graduation?*

## Method

### Participants

This study consisted of 44 respondents including 18 teacher educators, 16 student-teachers, and 10 mentor teachers (Table 1). In order to gain a more comprehensive picture of the components of the performance assessment scheme, cluster sampling was used to make the sample more representative of the teacher educators. Therefore, after identifying Farhangian centers that had applicants in English major, they were classified into five major sections based on their geographical regions. Then, the five sections were categorized further into smaller clusters on the basis of whether they had applicants in English major in 2012. Finally, the last stage included a simple random sampling from among the selected universities in each of the clusters. Therefore, the sample included six teacher educators from Tehran and three teacher educators from each of the following provinces: Khorasan, Eastern Azarbayjan, Khuzestan, and Mazandaran. Twelve teacher educators were selected on the basis of teaching practicum at the BA level whose applicants were accepted into university in 2012 and the remaining teacher educators were involved in this study mainly due to their experience and expertise in practicum. Their ages ranged between thirty-seven to sixty years old with the mean of 44.46. In terms of educational academic degree (Table 1), ten respondents had PhD degree in TEFL, five were PhD students studying TEFL, and three teacher educators had Master's degree. As regards the teaching experience of the participants, three teacher educators had 3-5 years of teaching experience and the

remaining ones had 23-33 years of teaching experience (Akbari & Dadvand, 2011; Gatbonton, 2008).

Table 1

*The Distribution of Participants Based on their Status and Degrees*

	B.A. student	B.A.	M.A.	PhD student	PhD	Total
Student- teacher	16	-	-	-	-	16
Mentor teacher	-	10	-	-	-	10
Teacher educator	-	-	3	5	10	18
total	16	10	3	5	10	44

The second group of participants that was used to elicit the components of performance assessment scheme was student-teachers studying B.A. level at Farhangian centers. Due to administrative constraints, a convenient sampling procedure was used to collect data from student-teachers. However, to make the sample more representative, attempts were made to conduct interviews with both male and female student-teachers in Tehran province. Hence, eight male student-teachers and eight female student-teachers were interviewed. Their ages ranged between 22 and 27 with the mean of 23.8. They were mainly selected from among those student-teachers who had passed practicum four.

This study also included ten teachers who taught at schools. Purposive sampling procedure was used to select the respondents from schools located in Tehran province. They were chosen from among the teachers who mentored student-teachers during their practicum four. They were all B.A. holders with 15-28 years of teaching experience. They fell in the age range of 30 to 48 with the mean of 37.

To collect the quantitative data, the developed scheme was sent to 200 respondents. Since access to the participants was not possible through other types of sampling procedures, convenience sampling was



used in this phase of the study (Ary, Jacobs, Sorensen, & Walker, 2014). Hence, majority of the participants were selected from Tehran province. The sample included 45 student-teachers, 63 teacher educators, 30 EFL teachers, 37 PhD students, and 25 university professors. From the total sample, 92 (46%) were males and 108 (54%) were females.

In terms of educational degree, 55 participants had PhD degree in TEFL, 54 participants were PhD students, 46 were MA holders, and 45 participants were student-teachers in the final year of university (Table 2). As regards the teaching experience of the participants, 52 respondents fell in the range of 0-4, 79 respondents were in the range of 5-12, 54 respondents fell in the range of 13-20, and 15 respondents had teaching experience of more than 20 years.

Table 2

*The Distribution of Participants Based on their Educational Degrees*

	B.A. student	M.A	PhD student	PhD	total
Student-teachers	45	0	0	0	45
EFL teachers	0	30	0	0	30
Teacher educators	0	16	17	30	63
PhD students	0	0	37	0	37
University professors	0	0	0	25	25
Total	45	46	54	55	200

**Instrument and Procedures**

The following instruments were used in this study:

**Stakeholders' interview guides.** In an attempt to identify the defining characteristics of a competent teacher in terms of practicum one, two, three, and four, in-depth interviews were conducted with the participants. Hence, the researchers were granted an official permission to conduct face-to-face interviews with the participants of the study at Farhangian centers. The aim of these interviews was two-fold: (a) to

identify the components of performance assessment from the perspectives of beneficiaries; and (b) to develop a valid instrument for measuring teacher performance. To this end, the researchers used the following sources to design interview questions:

- The National Curriculum Document of Tarbiat Moalem University
- The Curriculum Document of the English major at Tarbiat Moalem University

The interview guide (Appendix A) included ten questions. The clarity of the questions was checked by an expert in testing and also an expert in teacher education. Each interview lasted 30-50 minutes on average and all the interviews were recorded and transcribed by one of the researchers.

**Performance Assessment Scheme (PAS).** A performance assessment scheme was constructed and distributed among 200 participants. To develop the performance assessment scheme, the researchers triangulated data obtained from three sources: the content analysis of interviews, the content analysis of the Curriculum Document of the English major (the practicum part), and literature review. Using Danielson's framework as a guide, first, they extracted the themes from the study of literature, the content analyses of the interviews, and the Curriculum Document and categorized them. In a further analysis, the researchers left out the redundant themes and compiled the first draft of a list comprising the components of performance assessment. One item was added to the scheme from the analysis of the Curriculum Document of the English major (item 17), two items were exclusively extracted from interviews (items 12, and 13) and the remaining items were taken from Danielson's (2011) framework. On the whole, the scheme consisted of seventeen items in three main categories with five-point Likert-type scale: 1. *Planning and Instruction* (ten items); 2. *The Classroom Environment* (three items); 3. *Professional Development and Responsibilities* (four items) (Appendix B).

### Data Analysis

**Stakeholders' interview guides.** The data obtained in this phase of study were analyzed based on a qualitative-quantitative method. Following Miles and Huberman's (1994) suggestion, the analyses of the contents of interviews were carried out in two separate phases. First, a "vertical analysis" (Miles & Huberman, 1994) was done, in which the contents of interviews with each of the participants were analyzed separately from each other. Then, the second phase of the analysis called "horizontal analysis" (Miles & Huberman, 1994), with the intention of identifying common patterns and differences, was conducted. What follows exemplifies how each of the interviews was analyzed.

In horizontal analysis, first, the transcripts of interviews were analyzed using 'inductive analysis procedure' (Corbin & Strauss, 1990; Nouri & Mehrmohammadi, 2011). To further facilitate the process of categorization, Danielson's (2011) framework for teaching was used as a guide. In order to identify thematic categories, the transcribed interviews were segmented into chunks, each with a distinct theme. Then, the chunks were categorized into meaning units in terms of their content. Finally, each meaning unit was classified under its relevant category. In a further analysis, each category was labeled. What follows is an excerpt from the transcripts of an interview conducted with a female student-teacher. It clarifies the segmentation, categorization, and labeling processes.

"We gained considerable experience from practicum four. We were expected to prepare our lesson plan in advance (1). However, when it came to implementing it, we faced serious problems, because there were some weak students (2). Then we had to think of a way (3) to help them (4)".

The above extract from a transcribed interview was segmented into four chunks. After examining the chunks in terms of their underlying theme, they were classified under their relevant categories. For instance, while the first and the second chunks belonged to the meaning units

"*Designing lesson plans and unit structure*" and "*Demonstrating knowledge of students*", which were later classified under the category "*Planning and preparation*", the third chunk represented the meaning unit "*Use of a variety of instructional strategies*" and the fourth one belonged to "*Engaging students in learning*" which were later categorized under the category "*Instruction*". In an attempt to check the interrater dependability of the analysis, a subset of the transcripts was coded by a second party, an EFL head teacher familiar with Danielson's framework (2011); then, Cohen's Kappa statistic was used to determine if there was agreement between the two raters (Cohen, 1960).. The resulting kappa value of .88 indicated a strong agreement.

**Performance assessment scheme.** Two EFL experts reviewed the items of the scheme and offered suggestions to improve the wording of some of the items. Accordingly, changes were made to the items. The resulting version of the PAS was pre-administered to 38 EFL teachers. The reliability of the PAS was acceptable as the Cronbach Alpha coefficient was 0.70. Finally, the last version of the scheme was sent to all of the participants of the second phase of study. The participants were asked to show their level of agreement with each of the items ranging from 1 (*strongly disagree*), 2 (*disagree*), 3 (*undecided*), 4 (*agree*), to 5 (*strongly agree*), and rate the importance of them ranging from 1 (*less important*), 2 (*moderately important*), to 3 (*important*). Participants' answers were subjected to Confirmatory Factor Analysis using SPSS version 18.

### Results and Discussion

As pointed out earlier, the present study aimed to investigate the components of performance assessment, and consequently develop a performance assessment scheme. To begin to present the results of the study, the Cronbach Alpha coefficient for the Performance Assessment Scheme (PAS) is presented in Table 3. The reliability of the scheme was acceptable as the result showed an index of 0.70.

Table 3

*Alpha Cronbach Coefficient*

Cronbach's Alpha	N of Items
.706	13

Prior to performing CFA, the suitability of the data for factor analysis was assessed using the Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy (version 18). The KMO value was .699, above the recommended value of .5, and Bartlett's test of sphericity reached statistical significance ( $\chi^2(136) = 318.238, p < .05$ ) (Table 4), supporting the factorability of the correlation matrix.

Table 4

*The results of KMO and Bartlett's Test*

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.699
Bartlett's Test of Sphericity	Approx. Chi-Square	318.238
	df	136
	Sig	.000

A principal component analysis using VARIMAX Rotation was conducted to determine the factor structure of the scheme. The PCA analysis revealed the presence of seven factors with eigenvalues exceeding 1, explaining 15.21%, 9.795%, 9.079%, 7.947%, 7.605%, 6.475%, and 6.101% of the total variance respectively (Table 5, see Appendix C). Therefore, almost 62% of the total variance was explained by the first seven factors. As shown by Table 6, the clustering of items into factors seemed easily interpretable. Due to the content analysis of interviews with teacher educators and the general outline used in teaching practicum courses, a three-factor solution was selected explaining a total of 34% of the variance, with factor one contributing 15.21%, factor two contributing 9.8%, and factor three contributing 9.1%. In addition to content relevance, a loading greater than .3 on all factors was accounted for item inclusion in each distinct factor. Thirteen items fulfilled this criterion. However, four items from the seventeen

items of the scheme did not have enough factor loadings, less than .3, on any of the factors and were excluded.

Table 6

*Estimate Factor Loadings (Item loadings)*

Items	Factor 1	Factor 2	Factor 3
Item2	.353	.	.
Item3	.554	.	.
Item5	.510	.	.
Item1	.316	.	.
Item9	.499	.	.
Item12	.593	.	.
Item11	.590	.	.
Item6	.	.445	.
Item7	.	.374	.
Item8	.	.401	.
Item14	.	.	.718
Item16	.	.	.541
Item17	.	.	.668

Extraction Method: Principal Component Analysis (PCA)

Detailed inspection of the factor loadings displayed that seven items loaded most strongly on the first factor (Table 6). Since the items all relate to designing and implementing, the researcher labeled this factor *Planning and Instruction*. This factor mainly deals with the skills and knowledge teachers must be equipped with in order to organize the content to be taught and the implementation of the plans. *Demonstrating knowledge of students* (item 1), *Designing learning activities* (item 2), *Developing instructional materials and resources* (item 3), *Designing student assessments* (item 5), *Time Management* (item 9), *Engaging students in learning* (item 11), and *Using assessment in instruction* (item 12) belong to the factor.

The content analysis of the next three items: *Creating an environment of respect and rapport* (item 6), *Managing student behavior* (item 7), and *Organizing physical space* (item 8) that loaded on factor

two revealed that they all referred to different aspects of classroom management. In fact, this factor is concerned with teachers' skills and competencies to create an environment conducive to learning (Danielson, 2011). Finally, the last group of items: *Reflecting on teaching* (item 14), *Growing and developing professionally* (item 16), and *Reflective observation* (item 17) that were kept in this model as a result of factor analysis loaded on factor three. The study of the literature and the analyses of interviews indicated that this factor, represented as *Professional Development and Responsibilities*, is associated with being a true professional educator; it encompasses the roles assumed outside of and in addition to those in the classroom with students" (Danielson, 2011, p. 30).

In a further step to validate the performance assessment scheme, Chi-Square Test was used to check whether there were significant differences in participants' responses regarding item weighting (Tables 7, see Appendix D). The results indicated that items 11, 14, 12, 17, and 16, which are *Engaging Students in Learning*, *Reflecting on Teaching*, *Using Assessment in Instruction*, *Reflective Observation*, and *Growing and Developing Professionally* respectively, were considered the most focused items by the participants of this study. Items 3, 7, and 8 which are *Developing Instructional Material and Resources*, *Managing Student Behavior*, and *Organizing Physical Space* respectively, were not considered as focused items. In general, except for two items, significant differences were observed for all the items.

A more detailed analysis of the results (Table 8) revealed that there were statistically significant differences in *Designing Learning Activities* (item 2) ( $\chi^2 = 19.2, p < .05$ ), *Developing Instructional Materials and Resources* (item 3) ( $\chi^2 = 15.1, p < .05$ ), *Designing Student Assessments* (item 5) ( $\chi^2 = 86.2, p < .05$ ), *Managing Student Behavior* (item 7) ( $\chi^2 = 11.56, p < .05$ ), *Organizing Physical Space* (item 8) ( $\chi^2 = 27.5, p < .05$ ), *Time Management* (item 9) ( $\chi^2 = 8.6, p < .05$ ), *Engaging Students in Learning* (item 11) ( $\chi^2 = 97.1, p < .05$ ), *Using Assessment in Instruction* (item 12) ( $\chi^2 = 53.3, p < .05$ ), *Reflecting on Teaching* (item 14) ( $\chi^2 =$

67.7,  $p < .05$ ), *Growing and Developing Professionally* (item 16) ( $\chi^2 = 90.4$ ,  $p < .05$ ), and *Reflective Observation* (item 17) ( $\chi^2 = 87.1$ ,  $p < .05$ ). Only two items, *Demonstrating Knowledge of Students* (item 1) ( $\chi^2 = 3.3$ ,  $p > .05$ ) and *Creating an Environment of Respect and Rapport* (item 6) ( $\chi^2 = 3.6$ ,  $p > .05$ ) did not show statistically significant differences (Table 7).

Teacher evaluation requires addressing some central issues such as what of assessment (Taut & Sun, 2014). Therefore, it brings to the forefront the need for a solid understanding of the task of teaching. This argument is in line with Isoré's (2009) assertion saying that a critical touchstone necessarily preceding teacher evaluation is the definition of the criteria and standards of good teaching in terms of competencies and responsibilities. In this respect, the researchers made an effort to make a list of the competencies that student-teachers must be equipped with after passing practicum. What follows is a brief explanation of the items that were considered as important criteria for assessing teacher performance.

The results show that *Engaging students in learning* (item 11) is the most statistically significant item from the participants' perspectives (Table 8). This finding is supported by extensive literature (e.g., Danielson, 2011; Ellett, 1990; Skowron, 2001) suggesting that engaging students in learning is usually aimed at providing students with opportunities to develop important concepts and skills. Skinner and Belmont (1993) refer to an array of effective teachers' behaviors including modeling, provision of choice, and sincere praise aimed at engaging students. Similarly, Brophy and Good (1974) note that effective teachers engage their students with an enthusiastic style of teaching. Danielson (2011) also points out that engaging students is both a main responsibility of teachers and a principle of effective teaching.

The second item that revealed highest statistical significance was *Reflecting on Teaching* (item 14) (Table 8). Reflecting on teaching has received increasing appeal in the context of teacher education programs. Many researchers believe that true professional teachers are characterized by their ability to reflect on their teaching (Danielson, 2011; Skowron,



2001). Reeves (2004) appears to have embraced the same view when he states that it is through the process of reflection that the popularity of teaching techniques and their effectiveness can be told apart. Zeichner and Liston (1987) also maintain that reflective teacher education enables teachers to analyze their practices and direct their professional growth and development of the educational environment .

The third item with the highest statistical significance is *Using Assessment in Instruction* (item 12) (Table 8). As an essential factor in providing superior educational opportunities, using assessment in instruction is a key characteristic of effective teaching (Skowron, 2001). Considering monitoring students as a type of *Using Assessment in Instruction*, Skowron (2001) believes that it is through monitoring students that teachers can obtain diagnostic feedback and therefore plan follow-up steps. In the same vein, Kyriacou (2007) assumes that effective teachers must be judged based on their skills to monitor students and provide corrective feedback. More specifically, Training and Development Agency for Schools (TDA) published a list of professional standards that teachers needed to be equipped with prior to obtaining Qualified Teacher Status (QTS). Assessment and monitoring were two of the professional skills proposed as a requirement for the QTS .

Item 17, *Reflective Observation*, has also been considered a statically significant item from the participants' perspectives (Table 8). This item has been added to the questionnaire based on the analysis of the Curriculum Document of the English major. The document refers to the critical role of *Reflective Observation* in developing the professional competencies of ELT student-teachers. It is assumed that investigating the educational and pedagogical problems through reflectively observing school and classroom paves the way for independent and professional teaching and equips student-teachers with first-hand experiences. Kyriacou (2007) holds that observation is valuable for its contribution to professional development and to the creative tension that stimulates novice teachers' ideas about their teaching.

*Growing and Developing Professionally* (item 16) is another item reported as statistically significant which is one of the professional responsibilities of teachers (Danielson, 2011) (Table 8). It results in the improvement of teachers' content knowledge and pedagogical skills (Danielson, 2011). Fullan (2001) also acknowledges the significance of *Growing and Developing Professionally* by referring to the need for schools to create professional learning communities. In a similar view, Dufour and Eaker (1998) state that school improvement can be achieved if school personnel work as professional learning communities.

*Designing Learning Activities* (item 2) was considered a statistically significant item too (Table 8). Danielson (2011) believes that good instructional planning is essential to student learning and goes on to explain that *Designing Learning Activities* as a main component of instructional planning is also critical to student achievement. Kyriacou (2007) argues that the essence of effective teaching lies in designing activities that result in the kind of learning that teacher's desire. It is assumed that good teaching takes place when a variety of learning activities are designed and employed (Ofsted, 2006).

Another statistically significant item was *Time Management* (item 9) (Table 8). Effective teachers are characterized by their efficient use of time. In fact, *Time Management* is a critical component of a well-managed classroom in that it helps students achieve learning goals and contributes to a better learning environment in which students have a more positive attitude toward learning (Clement, 2010; McLeod, 2003). Kyriacou (2007) believes that the ability to manage time must be a main criterion for teacher appraisal .

*Designing Student Assessment* (item 5) was also a significant item (Table 8). In this regard, McBer (2000) proposes that assessment is one of the main teaching skills and responsibilities. The Department for Education and Skills (DfES) (2004) has also referred to designing assessment as one of the teaching skills that underlie good classroom practice and underscores the significance of assessment by alluding to it as the defining characteristic of an outstanding lesson. Similarly,

Kyriacou (2007) claims that "the skills involved in assessing pupils' progress, covering both formative (i.e. intended to aid pupils' further development) and summative (i.e. providing a record of attainment) purposes of assessment" (p. 11) are among the essential teaching skills.

Although the results indicate that *Developing Instructional Materials and Resources* (item 3), *Managing Student Behavior* (item 7), and *Organizing Physical Space* (item 8) were statistically significant (Table 8), the participants did not consider the three above mentioned items as important. Based on the literature, *Managing Student Behavior* is a factor characterizing competent teachers (Jackson, Simoncini, & Davidson, 2013). In fact, highly-qualified teachers are recognized by the various strategies they employ to manage student behavior (Clement, 2010). Whitaker (2004) underscores the importance of managing student behavior in this way: "Great teachers are very clear about their approach to student behavior. They establish clear expectations at the start of the year and follow them consistently as the year progresses." (pp. 17–18). Similarly, McCormick and Shi (1999) consider the ability to manage student behavior as a main indicator of teacher's professional identity. It is widely accepted that the ability to manage student behavior must be taken into account when evaluating teacher performance (McKenzie, Rowley, Weldon & Murphy, 2011).

Moreover, in well-organized classrooms where seating arrangement are carefully designed, students are encouraged to participate in classroom activities (Edwards, 1993). McLeod (2003) acknowledges the importance of *Organizing Physical Space* in this way:

Teachers try to make every inch of classroom space count in order to have a rich and inviting classroom environment because they know that the richness of students' experiences are enhanced or diminished by their surroundings. The organization of space also affects the way students behave and move around the classroom, as well as how much attention they pay to instruction (p. 3).

Danielson (2011) believes that creating and interpreting materials demonstrates teachers' competencies in several other areas. For instance, she states that classroom assignment is indicative of teachers' ability to engage students in learning and claims that teachers are responsible for developing instructional materials. Marzano (2003) also cites developing classroom materials as a key teaching skill. In the list of essential teaching skills contributing to effective teaching proposed by Kyriacou (2007), preparing materials and resources are also considered critical to planning and preparation for teaching .

In-depth analyses of the reflective commentaries of the student-teachers and also the interviews reveal the logic behind this weighting. Although they consider these three items as important competencies to be acquired by teachers, they believe that the items should not be of importance in evaluating their performance for several reasons. For instance, they believe that the way they were introduced to the students by their mentor teachers could adversely affect their ability to manage student behavior. What follows is an excerpt from the reflective commentaries of one of the student-teachers:

When I entered the class, the mentor teacher turned to the students and said, "she is an apprentice who wants to learn how to teach. Please let her perform her show". I hated the way she introduced me to the students because all of them laughed at me and I realized that I could not control the students easily anymore. For instance, when later the same day, I asked one of the students sitting in the front row to change her seat with one of her classmates at the back of the class, she ignored my request. I was not considered an authority in the classroom because of what the mentor teacher had told about me. That is why students did not listen to me .

In further analyses of the reflective commentaries and the interviews, the researcher found that student-teachers assumed that the skill in *Developing Instructional Materials and Resources* is of critical importance. However, they believed that this factor should not be a

criterion for evaluating their competencies unless they receive practical instructions on how to develop materials.

Taken together, the research base offers good reasons to attend vigilantly to the role performance assessment plays in learners' academic achievement generally and in teacher candidates' professional development, specifically (Danielson & Marquez, 1998; Delandshere & Arens, 2003; Haertel, 1991). As suggested by Darling-Hammond (2010), performance assessment assists student-teachers in their professional development by increasing their knowledge of teaching elements including subject matter knowledge, classroom management, and instructional planning. The analysis of the reflective commentaries of the student-teachers showed that their awareness of the components of the assessment scheme had a positive influence on their perception of teaching and motivated them to change their teaching practices (Darling-Hammond, 2010). In other words, "increased attention to mastery of the explicit criteria would, in most cases, result in increased competence for those affected" (Cizek & Bunch, 2007, p. 8).

Examining the effects of performance assessment on pre-service teachers also offers insights into improving the curriculum and the programs that prepare candidates for such evaluations (Selvester, Summers, & Williams, 2006). For instance, the analysis of the follow-up interviews with student-teachers indicate that if Farhangian University intends to assess student-teachers against the items *Developing Instructional Materials and Resources* (item 3), *Managing Student Behavior* (item 7), and *Organizing Physical Space* (item 8), first, it should offer some practical courses in this regard. It follows that understanding the value of performance assessment in teacher education programs has important implications, both for rethinking the design of the curriculum and for the professional development of student-teachers.

### **Conclusion**

The increasing prevalence of performance assessment in teacher education programs have led to burgeoning research on exploring

different aspects of teacher performance and teaching quality. This study is another attempt to provide insights into performance assessment by exploring its components in terms of competencies student-teachers are expected to acquire from practicum. Isoré (2009) asserts that a clear statement of the criteria and standards of good teaching in terms of competencies and responsibilities is an essential prerequisite for any teacher evaluation scheme. In a similar argument concerning the necessity of a conception of standards to examine teacher performance, Cochran-Smith (2001) acknowledges that a detailed description of what teacher candidates are expected to accomplish in terms of knowledge, skills, and dispositions is critical to the development of a valid teacher assessment. Against this background, as discussed above, interviews were conducted with 44 stakeholders to explore the components of performance assessment from their perspectives. Then, data obtained from the content analyses of the interviews, review of literature, and content analysis of the Curriculum Document of the English major were triangulated. Finally, a performance assessment scheme with 17 items was developed which was reduced to 13 items after Factor Analysis to be used as a benchmark for assessing the professional competencies of Farhangian University's ELT student-teachers. The thirteen items of the scheme loaded on three factors. The first factor, *Planning and Instruction*, consisted of seven items and dealt with skills that were necessary for planning instruction and successfully implementing it. Three items loaded highly on the second factor which was labeled as *Classroom Environment* since the three items related to the competencies that are required for classroom management. The last factor, *Professional Development and Responsibilities*, included three items pertaining to the professional lives of the student-teachers.

However, the results of this study must be interpreted cautiously for several reasons. First, although the researchers followed Isoré's (2009) suggestions to define the standards of good teaching prior to developing the performance assessment scheme, due to difficulty with segmenting, coding, and categorizing the transcripts of interviews and content

analysis of the Curriculum Document, replication studies are required to substantiate the findings of the present study. Second, taking into account practicality issues, this study was restricted to the investigation of the competencies that student-teachers were expected to acquire from practicum.

### References

- Akbari, R., & Dadvand, B., (2011). Does formal education make a difference? A comparison of pedagogical thought units of BA vs. MA teachers. *Modern Language Journal*, 95(1), 44-60.
- Ary, D., Jacobs, L. C., Sorensen, C. K., & Walker, D. (2014). *Introduction to research in education*. Cengage Learning.
- Beare, H. (1989). The Australian policy context. *Teacher appraisal issues and approaches*. Australian Council for Educational Research.
- Berliner, D. C. (2005). The near impossibility of testing for teacher quality. *Journal of Teacher Education*, 56, 205–213.
- Brophy, J., & Good, T. (1974). *Teacher-student relationships: Causes and consequences*. New York: Holt, Rinehart and Winston.
- Bunch, G. C., Aguirre, J. M., & Téllez, K. (2009). Beyond the scores: Using candidate responses on high stakes performance assessment to inform teacher preparation for English learners. *Issues in Teacher Education*, 18(1), 103–128.
- Chung, R. R. (2007). Beyond the ZPD: When do beginning teachers learn from a high-stakes portfolio assessment? *Paper presented at the Annual Meeting of the American Educational Research Association*. Chicago, IL. Retrieved from [https://scale.stanford.edu/system/files/RRChungAERA2007\\_0.pdf](https://scale.stanford.edu/system/files/RRChungAERA2007_0.pdf)
- Cizek, G. J., & Bunch, M. B.(2007). *Standard setting: A guide to establishing and evaluating performance standards on tests*. UK. Thousand Oaks.
- Clement, M. C. (2010). Preparing teachers for classroom management: The teacher educator's role. *The Delta Kappa Gamma Bulletin*,

77(1), 41-44.

- Cochran-smith, M. (2010). Foreword. In M. M., Kennedy (Ed.), *Teacher assessment and the quest for teacher quality: A handbook* (pp. xiii-1). San Francisco: John Wiley & Sons.
- Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20(1), 37-46.
- Corbin, J., & Strauss, A. (1990). Grounded theory research: Procedures, canons and evaluative criteria. *Zeitschrift fur Soziologie, JG*, 19, 418-427.
- Danielson, C. (2011). *Enhancing professional practice: A framework for teaching*. Association for Supervision and Curriculum Development (ASCD).
- Danielson, C., & Marquez, E. (1998). *A collection of performance tasks and rubrics: High school mathematics*. Larchmont, NY: Eye on Education.
- Danielson, C., & McGreal, T. (2000). *Teacher evaluation to enhance professional practice*. Alexandria, Virginia: Association for Supervision and Curriculum Development (ASCD).
- Darling-Hammond, L. (2010). *Evaluating teacher effectiveness: How teacher performance assessments can measure and improve teaching*. Center for American Progress.
- Darling-Hammond, L., & Snyder, J. (2000). Authentic assessment of teaching in context. *Teaching and Teacher Education*, 16, 523-545.
- Delandshere, G., & Arens, S. A. (2003). Examining the quality of evidence in preservice teacher portfolios. *Journal of Teacher Education*, 54 (1), 57-73.
- Department for Education and Skills (DfES) (2004). *Removing barriers to achievement: the Government's strategy for SEN* (London, DfES Publications).
- DuFour, R., & Eaker, R. (1998). *Professional learning communities at work: Best practices for enhancing student achievement*. Solution Tree Press.



- Edwards, C. H. (1993). *Classroom discipline and management*. New York: Macmillan Publishing.
- Ellett, C. (1990). *A new generation of classroom-based assessments of teaching and learning: Concepts, issues and controversies from pilots of the Louisiana STAR*. Baton Rouge: College of Education, Louisiana State University.
- Fullan, M. (2001). *Leading in a culture of change*. New York: Jossey-Bass.
- Gasper, M. F., & Vieira, C. R., (2013). Plentitude teacher evaluation for teacher effectiveness and well-being with neuro-linguistic programming. *US-China Education Review*, 3(1), 1-17.
- Gass, S. M., & Mackey, A. (2000). *Stimulated recall methodology in second language acquisition*. Mahwah, NJ: Erlbaum.
- Gatbonton, E. (2008). Looking beyond teachers' classroom behaviour: Novice and experienced ESL teachers' pedagogical knowledge. *Language Teaching Research*, 12(2), 161–182.
- Hatch, E. M., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. New York, NY: Newbury House Publishers.
- Haertel, E. H. (1991). New forms of teacher assessment. *Review of Research in Education*, 17, 3–29.
- Hamilton, L. (2003). Assessment as a policy tool. *Review of Research in Education*, 27, 25–68.
- Hodge, K. J., & Morgan, G. B. (2014). The stability of teacher performance and effectiveness: Implications for policies concerning teacher evaluation. *Education Policy Analysis Archive*, 22.
- Isoré, M. (2009), Teacher evaluation: Current practices in OECD countries and a literature review, *OECD Education Working Papers*, No. 23, OECD Publishing.
- Jackson, C., Simoncini, K., & Davidson, M. (2013). Classroom profiling training: Increasing preservice teachers' confidence and knowledge

- of classroom management skills. *Australian Journal of Teacher Education*, 38(8), 30-46.
- Koirala, H. P., Davis, M., & Johnson, p. (2008). Development of a performance assessment task and rubric to measure prospective secondary school mathematics teachers' pedagogical content knowledge and skills. *Journal of Math Teacher Education*, 11, 127-138.
- Kyriacou, C. (2007). *Essential teaching skills*. London, England: Nelson Thornes Ltd.
- Marzano, R. J. (2003). *What works in schools: Translating research into action*. Association for Supervision and Curriculum Development (ASCD).
- McBer, H. (2000). *Research into teacher effectiveness. A model of teacher effectiveness* (Vol. 216). Research report.
- McCormick, J & Shi, G. (1999). Teachers' attributions of responsibility for their occupational stress in the people's Republic of China and Australia. *British Journal of Educational Psychology*, 69(3), 393-407.
- McGreal, T. M. (1988). Evaluation for enhancing instruction: Linking teacher evaluation and staff development. In Stanley, S. J., & Popham, W. J. (Ed.). *Teacher evaluation: Six prescriptions for success*. Alexandria, Virginia: Association for Supervision and Curriculum Development.
- McKenzie, P., Rowley, G., Weldon, P., & Murphy, M. (2011). *Staff in Australia's schools 2010: Report prepared for the department of education, employment and workplace relations*. Melbourne: Australian Council for Educational Research (ACER).
- McLeod, J. (2003). Time and classroom space. In J. McLeod, J. Fisher, & G. Hoover. *The key elements of classroom management. Managing time and space, student behavior, and instructional strategies*. Association for supervision and curriculum development

- Medley, D. M. (1982). *Teacher competency testing and the teacher educators*. Association of Teacher Educators and the Bureau of Educational Research, University of Virginia.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: A sourcebook*. Beverly Hills: Sage Publications.
- Mitchell, K. J., Robinson, D. Z., Plake, B.S., & Knowles, K.T. (2001). *Testing teacher candidates, the role of licensure tests in improving teacher quality*. National Academy Press. Washington, D.C.
- Nouri, A. & Mehrmohammadi, M. (2011). A model for applying the grounded theory in educational research. *Journal of Curriculum Studies*, 23, 8-35.
- Navidinia, H., Kiani, G. R., Akbari, R., & Ghaffar Samar, R. (2015). EFL teacher performance evaluation in Iranian high schools: Examining the effectiveness of the status quo and setting the Groundwork for developing an alternative model. *The International Journal of Humanities*, 21(4), 27-53.
- Ofsted (2006). *The annual report of her majesty's chief inspector of schools 2005/2006*. London: The Stationery Office.
- Pecheone, R., & Chung, R.R. (2006). Evidence in teacher education: The performance assessment for California teachers. *Journal of Teacher Education*, 57(1), 22-36.
- Peterson, K. D. (2000). *Teacher evaluation: A comprehensive guide to new directions and practices* (2<sup>nd</sup> Ed.), Thousand Oaks, CA: Corwin Press.
- Prybylo, D. (1998). Beyond a positivistic approach to teacher evaluation. *Journal of school Leadership*, 8, 558-583.
- Reeves, D. B. (2004). *Accountability for learning: How teachers and school leaders can take charge*. Alexandria, VA: Association for Supervision and Curriculum Development.
- Sanders, W. L., Wright, S. P., & Horn, S. P. (1997). Teacher and classroom context effects on student achievement: Implications for teacher evaluation. *Journal of personnel evaluation in education*, 11(1), 57-67.

- Sandholtz, J. H. (2012). Predictions and performance on the PACT teaching event: Case studies of high and low performers. *Teacher Education Quarterly*, 39(3), 103-126.
- Selvester, M. P., Summers, D., & Williams, F. E. (2006). Costs and benefits of accountability: A case study of credential candidates' performance assessment. *Issues in Teacher Education*, 15(1), 21-36.
- Shinkfield, A. J., & Stufflebeam, D. L. (1995). School professionals' guide to improving teacher evaluation systems. In *Teacher Evaluation* (pp. 81-172). Springer Netherlands.
- Shulman, L. (1986). Those who understand: knowledge growth in teaching. *Educational Researcher*, 15(4), 4-14.
- Shulman, L. (1987). Knowledge and teaching: Foundations of the new reform. *Harvard Educational Review*, 57(1), 1-22.
- Skinner, E. A., & Belmont, M. J. (1993). Motivation in the classroom: Reciprocal effects of teacher behavior and student engagement across the school year. *Journal of Educational Psychology*, 85(4), 571.
- Skowron, J. (2001). *Powerful lesson planning models: The art of 1,000 decisions*. Arlington Heights, IL: Sky Light Training and Publishing.
- Taut, S., & Sun, Y. (2014). The development and implementation of a national, standards-based, multi-method teacher performance assessment system in Chile. *Education Policy Analysis Archives*, 22 (71).
- Torgerson, C. W., Macy, S. R., Beare, P., & Tanner, D. E. (2009). Fresno assessment of student teachers: A teacher performance assessment that informs practice. *Issues in Teacher Education*, 18(1).
- Whitaker, T. (2004). *What great teachers do differently: Fourteen things that matter most*. Larchmont, NY: Eye on Education.
- Zeichner, K. M., & Liston, D. P. (1987). Teaching student teachers to reflect. *Harvard Educational Review*, 57(2), 3-48.

## **Appendix A**

### **Interview Guide of Phase I**

Following Farhangian University's recent policy on requiring student-teachers to pass a comprehensive exam prior to obtaining teaching licensure, this study is concerned with performance assessment as part of the requirement. This phase of study aims at identifying the components of the performance assessment scheme from the stakeholders' perspectives. Your candid responses will contribute to the success of this research effort.

1. What are the defining characteristics of an effective teacher?
2. Imagine you want to assess the professional competencies of a teacher and then based on his/her performance give him/her teaching credential. What are your criteria for such an assessment?
3. What core qualities should a student-teacher have after passing practicum one?
4. What should be the components of performance assessment based on the competencies student-teachers obtain from practicum one?
5. What core qualities should a student-teacher have after passing practicum two?
6. What should be the components of performance assessment based on the competencies student-teachers obtain from practicum two?
7. What core qualities should a student-teacher have after passing practicum three?
8. What should be the components of performance assessment based on the competencies student-teachers obtain from practicum three?
9. What core qualities should a student-teacher have after passing practicum four?
10. What should be the components of performance assessment based on the competencies student-teachers obtain from practicum four?

**Appendix B**  
**The Performance Assessment Scheme (PAS)**

	Components of performance assessment	Your Degree of Agreement					Weight of Components		
		SD 1	D 2	UD 3	A 4	SA 5	Less important 1	Moderately Important 2	Important 3
<b>Domain 1: Planning and instruction</b>									
1	Demonstrating knowledge of students' needs, goals, proficiency								
2	Designing learning activities								
3	Developing instructional materials and resources								
*4	Designing lesson plans and unit structure								
5	Designing student assessments (criteria, formative assessments)								
9	Time management								
11	Engaging students in learning								
12	Using assessment in instruction (feedback, monitoring...)								
*13	Using relevant teaching methods								
*10	Use of a variety of instructional strategies								
<b>Domain 2: The classroom environment</b>									



## Appendix C

Table 5: Results of total variance explained by factors

Component	Initial Eigenvalues			Extraction Sums of Squared Loadings			Rotation Sums of Squared Loadings		
	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %	Total	% of Variance	Cumulative %
	1	2.586	15.214	15.214	2.586	15.214	15.214	2.221	13.063
2	1.665	9.795	25.009	1.665	9.795	25.009	1.809	10.640	23.703
3	1.543	9.079	34.088	1.543	9.079	34.088	1.765	10.385	34.088
4	1.351	7.947	42.035						
5	1.293	7.605	49.640						
6	1.101	6.475	56.115						
7	1.037	6.101	62.216						
8	.933	5.486	67.702						
9	.899	5.286	72.988						
10	.786	4.624	77.612						
11	.724	4.259	81.871						
12	.656	3.857	85.727						
13	.565	3.323	89.050						
14	.536	3.154	92.204						
15	.472	2.775	94.979						
16	.461	2.711	97.690						
17	.393	2.310	100.000						

Extraction Method: Principal Component Analysis (PCA).



## Appendix D

Table 7: Results of  $X^2$  statistical test for item weighting

Items	Observed/Expected Freq	Weights			Chi- Square	Asym p.Sig.
		Less important	Moderately important	important		
Item1	Observed	53 (26.7%)	70(34.7%)	77 (38.7%)	3.3	.100
	Expected	66.6	66.6	66.6		
Item2	Observed	33 (16.7%)	79(39.3%)	88(44%)	19.2	.001
	Expected	66.6	66.6	66.6		
Item3	Observed	89 (44.5%)	76(38%)	35 (17.5%)	15.1	.000
	Expected	66.6	66.6	66.6		
Item5	Observed	60(30%)	81(40.5%)	59 (29.5%)	86.2	.021
	Expected	66.6	66.6	66.6		
Item6	Observed	43 (21.5%)	85(42.5%)	72(36%)	3.6	.301
	Expected	66.6	66.6	66.6		
Item7	Observed	69 (44.5%)	67(33.5%)	44(22%)	11.56	.002
	Expected	66.6	66.6	66.6		
Item8	Observed	104(52%)	61(30.5%)	35 (17.5%)	27.5	.000
	Expected	66.6	66.6	66.6		
Item9	Observed	45(22.5%)	71(35.5%)	84(42%)	8.6	.000
	Expected	66.6	66.6	66.6		
Item11	Observed	5(2.5%)	59(29.5%)	136 (68%)	97.1	.013
	Expected	66.6	66.6	66.6		
Item12	Observed	19(9.5%)	65(32.5%)	116(58%)	53.3	.000
	Expected	66.6	66.6	66.6		
Item14	Observed	15(7.5%)	61(30.5%)	124(62%)	67.7	.004
	Expected	66.6	66.6	66.6		
Item16	Observed	26(13%)	77(38.5%)	97(48.5%)	90.4	.000
	Expected	50.0	50.0	50.0		
Item17	Observed	55 (27.5%)	43(21.5%)	102 (51%)	87.1	.000
	Expected	50.0	50.0	50.0		

**Appendix E****Table 8: Statistically significant items**

order	Item	Weight	Chi-Square
1	Engaging students in learning (11)	136	97.1
2	Reflecting on teaching (14)	124	67.7
3	Using assessment in instruction (12)	116	53.3
4	Reflective observation (17)	102	87.1
5	Growing and developing professionally (16)	97	90.4
6	Designing learning activities (2)	88	19.2
7	Time management (9)	84	8.6
8	Designing student assessment (5)	59	86.2
9	Managing student behavior (7)	44	11.56
10	Developing instructional materials and resources (3)	35	15.1
11	Organizing physical space (8)	35	27.5