

Evaluation of In-service Teacher Training Program in Iran: Focus on the Kirkpatrick Model

Maryam Mahmoodi¹, Mojgan Rashtchi², Gholam-Reza Abbasian³

¹*Department of Foreign Languages, Faculty of Literature, Humanities, and Social Sciences,
Science and Research Branch, Islamic Azad University, Tehran, Iran,*

E-mail: maryammahmoudi75@yahoo.com

ORCID: <https://orcid.org/0000-0002-9013-607X>

² *(Corresponding editor) TEFL Department, Faculty of Foreign Languages, North Tehran Branch,
Islamic Azad University, Tehran, Iran,*

E-mail: mojgan.rashtchi@gmail.com

ORCID: <https://orcid.org/0000-0001-7713-9316>

³ *Imam Ali University, Tehran, Iran,*

E-mail: gabbasian@gmail.com

ORCID: <https://orcid.org/0000-0003-1507-1736>

DOI: 10.26907/esd14.4.03

Abstract

This study investigated the effect of in-service education and training (INSET) courses on grade 11 EFL teachers' knowledge base employing Kirkpatrick's four-level (reaction, learning, behavior, result) evaluation model (Kirkpatrick & Kirkpatrick, 2006). To this end, the researchers designed a study at four stages. At the first stage, a standardized questionnaire, English language teachers' knowledge base (ELTKB), was administered to thirty 11th grade teachers to examine their views regarding the INSET classes (reaction). Then ten teachers were interviewed, and their classes were observed to explore to what extent the INSET courses had affected teachers' perceptions and actual job performance (learning). In the next step, 126 students filled in the students' questionnaire (SQ) regarding their perception of their teachers' teaching performance in the English classes before and after their attendance in INSET classes. At the final stage, the performance of the 126 students on a language proficiency test was measured to study to what extent their teachers' performance had affected their achievement. The result of the ELTKB questionnaire, SQ, interviews, and observations revealed the beneficial role of INSET classes in teachers' *reaction, learning, and behavior*. Moreover, the *result* of the program was positive since grade 11 students' language learning improved after their teachers' INSET class attendance. This program evaluation can be illuminating for stakeholders, policymakers, and curriculum designers.

Keywords: EFL teachers, INSET courses, Kirkpatrick, program evaluation

Оценка курса повышения квалификации иранских учителей с использованием модели Киркпатрика

Марьям Махмуди¹, Мойган Рашчи², Голам-Реза Аббасьян³

¹ Исламский университет Азад, Тегеран, Иран

E-mail: maryamtahtmoudi75@yahoo.com

ORCID: <https://orcid.org/0000-0002-9013-607X>

² Исламский университет Азад, Тегеран, Иран

E-mail: mojgan.rashtchi@gmail.com

ORCID: <https://orcid.org/0000-0001-7713-9316>

³ Университет Имама Али, Тегеран, Иран

E-mail: gabbasian@gmail.com

ORCID: <https://orcid.org/0000-0003-1507-1736>

DOI: 10.26907/esd14.4.03

Аннотация

Работа авторов посвящена изучению того, как влияет курс повышения квалификации (КПК) на учителей 11-ых классов, с использованием четырехступенчатой модели Киркпатрика (реакция, обучение, поведение, результат). Для этого было проведено исследование в 4 этапа. На первом этапе стандартизированный опрос (оценка знаний учителей английского языка) был проведен среди 30 учителей 11-ых классов с целью определить их мнение о курсе обучения (реакция). Затем было проанализировано влияние КПК на действия учителей (обучение) при помощи десяти интервью с учителями и наблюдений в классах. Отвечая на вопросы анкеты, 126 учеников поделились мнением о качестве работы их учителей до и после КПК. На последнем этапе знания 126 учеников по английскому языку были оценены с помощью теста, что позволило понять, насколько изменения в процессе преподавания повлияли на их успеваемость. Результаты опроса, интервью и наблюдений выявили, что КПК оказал положительное воздействие на реакцию, обучение и поведение учителей. Дополнительным положительным результатом можно считать то, что уровень английского языка среди одиннадцатиклассников повысился после участия их учителей в КПК. Оценка учебной программы может быть полезна лицам, определяющим образовательную политику, и разработчикам учебных программ.

Ключевые слова: учителя английского языка, курс повышения квалификации, модель Киркпатрика, оценка программы.

Introduction

Teachers play a crucial role in any education system, and their well-preparedness can affect students' learning. Teachers' professional growth, which, in the long run, leads to the expansion of knowledge and understanding, is the goal of educators, policymakers, and administrators who work for the betterment of society. Therefore, evaluating the appropriateness of teacher training programs and the extent to which such programs can help teachers in their intricate career highlight the importance of studies on program evaluation.

In 2013, after three decades of employing methods which followed the fundamental characteristics of traditional teaching (i.e., Grammar-Translation, Reading, and Audio-Lingual), English language teaching in Iran witnessed a paradigm shift. The change in the teaching approach was reflected in the newly developed course books for three junior high school levels (published in 2013) and three high school levels (published in 2017). The new English textbooks are adjusted to Communicative Language Teaching (CLT)

approach by the National Curriculum Development Center. This transformation in the English language curriculum makes the concept of teacher training a central concern of foreign language teaching (Freeman, 2002) and has to be accompanied by a modification in teachers' mental approach. Employing a new approach to teaching, as Roberts (1998) rightly claims, "is not a simple matter of replacing materials, or expecting to train teachers to change their practice from 'old' to 'new' like respraying a car" (p. 272). Teachers need to change their perspectives and concepts toward learning and learners (Cohen & Hill, 1997; Grosso de Leon, 2001); they should be able to reflect the alterations in their way of teaching. They should adopt new strategies and techniques while expanding their knowledge base to meet the requirements brought about by the new teaching perspectives.

In-service education and training (INSET) courses are an excellent source of training for the changes that curriculum developers expect from teachers. The purpose of the courses is to help teachers to go through all the necessary professional developments and promote their teaching efficacy. The present study aimed to examine the effect of INSET courses on Iranian English teachers' actual job performance. From among different models, which are used for program evaluation, the researchers selected the Kirkpatrick (1998) model and followed the Kirkpatrick and Kirkpatrick (2006) guidelines for program evaluation. Donald Kirkpatrick first introduced this model in 1959, and since then it has been shown to be useful (Bates & Coyne, 2005; Gill & Sharma, 2013;) and has proven to be one of the influential models despite criticisms against its application (e.g., Bates, 2004; Spitzer & Conway, 2002; Swanson, 2001). One advantage of the model is that at each level, it enables researchers to triangulate their investigation by employing different data collection instruments. Additionally, the researchers do not need to measure or explain the variables that underlie and interact with the training process (Bates, 2004). Thus, the researchers of the current study selected the Kirkpatrick model to examine the efficacy of INSET courses for 11th-grade high school teachers.

The model has four levels of reaction, learning, behavior, and result. Reaction relates to participants' attitudes toward the program that they attend and includes components such as the topic, instructor, materials, activities, instruments, and the like. The second level, learning, aims to determine what the trainees have learned from the training program. Behavior, the third level of evaluation, attempts to discover whether the learned issues are applied in the workplace. The last level, result, inspects to what extent the training can be seen in trainees' performances and learners' achievements and is the most challenging level as it deals with the outcome of teaching (Brewer, 2011; Grosso de Leon, 2001).

Most of the previous studies on the efficacy of the teacher education courses in Iran, though instructive, have focused mainly on the reaction level of the model and have pried into teachers' perceptions and the degree of their satisfaction with the materials provided in the courses (e.g., Birjandi & Derakhshan Hesari, 2010; Egtesadi & Hassanabadi, 2016; Hashemian & Azadi, 2014; Kazemi & Ashrafi, 2014). Although Khanjani, Vahdany, and Jafarigohar (2016) have utilized Kirkpatrick's model for evaluating pre-service teachers' pedagogical content knowledge, their study does not consider learners' improvement. Thus, the present study, employing the four levels of Kirkpatrick's model, aimed to evaluate not only teachers' perceptions and their actual performance in the classroom but also focused on learners' change of behavior as one of the factors, which can reflect the success of a program and which has not been tackled by other researchers in Iran.

It is worth mentioning that although this study intended to shed light on the efficacy of INSET classes and their results on teachers' classroom performance, it was not flawless. The first limitation involved the selection of participants from one province in Iran. Second, this study only considered grade 11 teachers' perceptions. Additionally, using

other data gathering tools such as think-aloud protocols while teachers are involved in learning during INSET courses could provide more valuable information.

Literature Review

Program evaluation is a “systematic assessment of the operation and/or outcomes of a program or policy, compared to explicit or implicit standards, as a means of contributing to the improvement of the program or policy” (Weiss, 1998, p. 18). Thorough program evaluation in the domain of education should include the performances of students, teachers, and institutions (Kiely & Rea-Dickins, 2009). According to Kirkpatrick and Kirkpatrick (2006), improving a program, increasing the transfer of learning to behavior, and then to results, and demonstrating the value of training to the organizations are the three primary reasons for evaluation of training. When programs go through a change, similar to that in Iran, evaluating their characteristics becomes a necessity, and thus scholars are encouraged to study their usefulness as continuous evaluation of an education program can guarantee its effectiveness. Program evaluation, as Kiely (2009, p. 99) argues, “is a form of enquiry which describes the achievements of a given programme, provides explanations for these, and sets out ways in which further development might be realized.” As Brewer (2011) argues, the purpose of program evaluations is to examine whether a program has achieved its objectives, has been able to cause improvement, and has been able to confirm the decisions made by stakeholders to change a program.

Program evaluation experts have developed different models, which set the criteria for measuring the degree of efficiency of training programs. Kirkpatrick’s (Kirkpatrick & Kirkpatrick, 2006) four-level model, selected in the present study, has been employed in different fields and by different researchers. For example, Bledsoe (1999) evaluated training courses in a computer company to explore correlations among the four levels of the Kirkpatrick model. She found that only reaction and behavior levels of the model correlated moderately and that evaluations conducted at level one do not guarantee the overall success of a training program. Wertz (2005) studied the effectiveness of the CLAD (cross-cultural, language, and academic development) training for a group of K-12 teachers. The data collected from different research tools showed positive responses at levels one, two, and three of the model. Also, Badu (2013) used the model to evaluate a group of university students’ learning of Initial Value and Boundary Condition Problems in mathematics, utilizing a questionnaire and a performance rubric. The study showed that the model was useful for the evaluation of the students’ learning. In another study, Yusoff et al. (2016) considered the first two levels of Kirkpatrick’s model and assessed the effect of an in-service teacher training program on school-based assessment and teachers’ learning. The researchers concluded that for the program to be useful in promoting teachers’ abilities to use school-based assessments, it should be improved and implemented for a more extended period.

The present study implemented Kirkpatrick’s model and aimed to examine to what extent INSET courses were successful in increasing the knowledge base of 11th grade English teachers. Thus, the researchers formulated the following research questions:

1. Do Iranian EFL teachers believe that INSET programs promote their knowledge base? (Reaction)
2. Can INSET classes change the actual job performance of Iranian EFL teachers? (Learning)
3. Do EFL students perceive that their teachers’ teaching performance can change after their participation in INSET classes? (Behavior)
4. Does EFL teachers’ attendance in INSET classes affect their students’ language proficiency? (Result)

Method

The researchers designed a study at four stages to meet the levels of the Kirkpatrick model. A quantitative phase was followed by two qualitative phases to triangulate the data. In the first stage, a survey was performed to scrutinize the reaction of the teachers via a questionnaire (level one). Then in the qualitative phases, ten teachers volunteered to participate in interview sessions and agreed to be observed while teaching (level two). In the third stage, the researchers conducted another survey to examine the EFL students' perceptions regarding behavior (level three) of their teachers before and after participation in the INSET courses. For obtaining information regarding the result level (level four), the students' language proficiency was measured before and after the EFL teachers' participation in INSET classes. The design of the study in the last stage was one group pretest-posttest design (Best & Kahn, 2006).

Participants

At the onset of the study, based on purposive sampling, thirty 11th grade teachers were selected because the textbook of this level was substituted with the newly developed one, and teachers were invited to participate in online INSET classes to be trained for teaching the new book. The participants completed a questionnaire before and after attending the courses to help the researchers evaluate their perception regarding the training courses. Twenty-nine teachers filled in the questionnaire, and one did not return it. Ten of the teachers agreed to take part in the interviews and accepted being observed.

Moreover, from among 285 grade 11 students (132 girls and 153 boys), 126 students who got 40 or above in a 60-item English test, developed based on their textbook, were selected through purposive sampling to participate in a survey study. The reason for setting 40 as the cut-off score was to incorporate the evaluation of students who had a higher level of proficiency. The same students took an English pretest and posttest to find the effect of their teachers' way of teaching on the enhancement of their English language ability.

Instrumentation and Data Collection Procedure

The researchers developed two questionnaires, English language teachers' knowledge base (ELTKB) questionnaire (Appendix A), to collect data regarding the teacher participants' reaction. The questionnaire consisted of two sections. The first section aimed to collect the respondents' demographic information. The second section consisted of 56 close-ended items on a 5-point Likert-type scale from 1 (strongly disagree) to 5 (strongly agree). This part of the questionnaire asked about teachers' Content Knowledge (CK, five questions), Technological Knowledge (TK, five questions), Technological Content Knowledge (TCK, seven questions), Pedagogical Knowledge (PK, 13 questions), Pedagogical Content Knowledge (PCK, 12 questions), Technological Pedagogical Knowledge (TPK, ten questions), and Technological Pedagogical and Content Knowledge (TPACK, 13 questions), respectively. The questionnaire examined the teachers' perceptions regarding the usefulness of the INSET course they had attended.

The researchers followed three procedures to determine the construct validity of the questionnaire. First, they conducted a panel discussion on the items of the initial draft. Second, they sought advice from ten professionals in the field of applied linguistics, survey design and statistics, and teacher training experts. Finally, they performed Exploratory Factor Analysis comprising the correlation matrix, communalities, total variance, component matrix, rotated matrix, component transformation matrix, and reproduced correlation. Cronbach's alpha computed to estimate the reliability of the questionnaire showed an acceptable reliability index ($r=0.87$).

Semi-structured interviews and observations were employed in the second stage. Three English language professors and two English language teacher trainers reviewed the interview questions (Appendix B). Based on their feedback, the researchers revised the wording of the items and reduced the number of questions. The interviews aimed to explore to what extent the teachers believed that the classes helped them improve their knowledge regarding the new book. The interviews (approximately 20 minutes for each individual) were recorded, after gaining permission from the interviewees, and were transcribed later. Qualitative content analysis, open, axial, and selective, suggested by Strauss and Corbin (1990), was utilized to categorize the data. The transcripts and the related data coding procedures were checked by an authority in the field of education and research to ensure their reliability.

Checklists (Appendix C) were employed during the observations. Experts in the field acknowledged the content of the checklist. The checklists were used before the main study in three similar English classes to ensure their practicality. The observations and interviews could reveal to what extent the teachers had learned from the training courses (level two of the model). The classes of ten 11th grade teachers were observed four times, twice before (in October 2017), and twice after (March 2018) attending the INSET classes. Each observation took about 45 minutes.

Another instrument was the student questionnaire (SQ), which contained 21 items on a 5-point Likert-type scale to gather the students' opinions about their teachers' change of behavior before and after training (Appendix D). To determine the content validity of the questionnaire, three English language teachers, two INSET program instructors, and one English language professor went through the questions before piloting them. Cronbach's alpha estimated in the pilot study ($r=0.94$) and actual study ($r=0.92$) revealed the reliability of SQ. The questionnaire was distributed among 11th-grade students 20 minutes before the end of their classes to discover their perceptions about their teachers' teaching behavior. Both the school principal and teachers had signed a letter of consent before distributing the questionnaire.

Additionally, the researchers developed a general language proficiency test based on the 10th-grade textbooks to select the students who were qualified to answer the SQ. This test had 60 items and contained vocabulary, grammar, and reading comprehension questions (each section 20 items). The researchers piloted the test before the study and computed its reliability through the KR-21 formula ($r=0.79$). This test was administered among 285 grade 11 students from among whom 126 were selected to participate in the third and fourth stages of the study.

The researchers used a further test to measure the language proficiency of the 126 students whose teachers had registered to take part in INSET classes. The purpose was to enable researchers to examine the effect of the teachers' knowledge base on students' language achievement and answer to Kirkpatrick's level four evaluation model (result). The students took a standardized test adopted from a test series published by a university entrance preparation center at the end of November 2017 when still the teachers were taking part in INSET classes. The same test was administered as the posttest at the end of March 2018. Two university professors reviewed the test before piloting it with 30 students similar to the participants, and Cronbach's alpha showed an acceptable reliability index ($r=0.82$).

Results

Grade 11 English Teachers' Perception

A paired samples t-test was run to answer the first research question. Table 1 shows an increase in the mean of the teachers' perception before ($M=2.25$, $SD=0.16$) and after ($M=3.13$, $SD=.17$) attending the INSET training courses.

Table 1. Descriptive statistics, teachers' perception before & after INSET classes

	Mean	N	SD	Std. Error Mean
Before	2.2557	29	.16841	.03127
After	3.1338	29	.17055	.03167

The result of the paired sample t-test (Table 2) shows a statistically significant difference in the teachers' perception before and after attending the INSET courses ($t(28) = -28.80, p \leq 0.01$). Thus, the teachers' reaction (level one of the Kirkpatrick model) was positive toward the classes. The mean increase in teachers' perception was .87 with a 98% confidence interval ranging from -0.94 to -0.85. The η^2 statistic (0.96) showed a large effect size, indicating that 96% of the change in the teachers' perception was related to the training they had received.

Table 2. Paired sample T-test, teachers' perception before & after INSET class attendance

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	SD	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Before & after INSET	.87810	.16417	.03049	-.94055	-.81565	-28.803	28	.000

Interviews

The results of the interviews are categorized according to different components of teachers' knowledge base to facilitate answering to the second research question. Regarding CK, the interviewees attributed a broader meaning to content knowledge and asserted that the classes should give them knowledge about the scientific issues discussed in the book. They believed that having such knowledge could enable teachers to be involved in a conversation with the students. However, despite complaints, they stated that they had gained knowledge about the theoretical and practical characteristics of the CLT approach. One issue was that teachers expected more emphasis on the four language skills (reading, writing, speaking, listening) and sub-skills (grammar, pronunciation, vocabulary).

Additionally, they stated that culture, a feature absent in INSET classes, could promote teachers' knowledge base and must be an inseparable part of language teaching in the future courses. They believed a focus on culture could facilitate teaching for teachers and learning for learners. For example, explicit explanation about differences in greeting, apologizing, and requesting between Persian and the English could make the classes more fruitful.

Regarding TK, the interviewees complained about the quality of the technological classes since the classes were instructed by teachers who were efficient in technology but not in English. They believed that the trainees could have benefited from more exposure to English if their teachers had been able to speak English. Overall, the teachers were not completely satisfied with the quality of the technology courses though they believed that they had learned many things from the classes.

As far as TCK was concerned, the teachers believed the classes could help them build a community through social media such as Instagram and Telegram to share their teaching and learning experiences with their colleagues. The disadvantage of the classes, as they asserted, was that the objective of the classes was not clear. The teachers expected to learn more about the application of technology in their classes and increase their knowledge

base. PK received more attention than the other knowledge types, and the interviewees reported that they had learned a lot since the INSET classes could provide relevant insights and knowledge about the practical ways of student assessment, classroom management, lesson planning, and teaching methodologies and strategies. However, from the teachers' viewpoints, the weak aspect of the classes was the lack of concentration on reflective teaching.

The interviewees' responses to the role of the classes in enhancing their PCK could be categorized into two viewpoints. Five of the teachers believed that the classes guided them to prepare lesson plans and to engage the students in speaking rather than merely in grammar and vocabulary. However, five teachers criticized the classes for not relating their CK to PK. Regarding TPK, the teachers unanimously criticized the classes for not presenting the lessons via technology. They complained about not being capable of integrating technological and pedagogical knowledge. They believed that the classes should be more practical and emphasized on the role of workshops in improving TPK. Also, they suggested the use of films, slides, real-life tasks in INSET classes.

Four of the interviewees believed that TPACK in INSET classes did not prepare them to use different language teaching strategies via technology. However, six of the teachers found the classes useful and stated that they could employ technology to make their classes an active environment for learning. They stated that they had learned how to use PowerPoint, and search the Internet to find appropriate pictures and cartoons related to the lessons. For them, learning to use technology had motivated them to teach meaningfully. However, the interviewees felt that the duration and content of the classes were not sufficient, and underscored the role of actual practice under the supervision of instructors. All teachers expected authorities to equip schools with technological devices as they could contribute to successful teaching.

Teacher Observations

The first two observations took place before the teachers' INSET class attendances at the onset of the school year. The teachers were involved in teaching the first lesson of the newly published book on CLT. The lesson began with the languages of the world, and the students were asked to guess the language spoken in each of the countries on the map. According to the observation field notes, almost all teachers followed the same procedures. All classes were teacher-centered with no group work activities. The teachers were mostly content-driven. Traces of ALM (repetition and memorization) were observable in the classes, and the medium of instruction was students' L1. Teachers were reluctant to spend time and energy on speaking and writing skills. Only two of the teachers were familiar with the CLT approach. Students' errors were corrected immediately and explicitly. Regarding technological devices, only two classes out of ten had access to a DVD player and a CD player, and two of the teachers used their cell phones to play the listening excerpt.

The second two observations occurred six months after the first ones when the teachers were teaching lesson three, the final lesson of the book. The observer, who was one of the researchers, witnessed more brainstorming activities and well-organized teaching schedules. The teachers taught listening and reading comprehension strategies to facilitate learning and were more tolerant of students' errors. The teachers guided students' interactions and engaged them in classroom activities. The observer reported the dominance of CLT and the use of meaningful activities, although students' inadequate level of English proficiency made them teach grammar deductively.

The teachers' inclination to incorporate technology into their teaching practice also had increased. For example, they used a speaking dictionary to check the pronunciation of the words. One of the teachers encouraged students to look for supplementary materials

from the Internet and introduced some websites. However, the lack of technological support hindered them from establishing a relationship between technology and teaching. Besides, teachers' lack of time and large class size had reduced the quality of teaching. In summary, considering the interviews and classroom observations, researchers believe that the INSET courses could boost the teachers' knowledge base, and answer the second research question affirmatively.

Students' Perceptions Regarding Teachers' Performances

Table 3 shows the descriptive statistics of the students' questionnaires before ($M=2.48$, $SD=0.74$) and after ($M=3.24$, $SD=1.08$) their teachers' INSET class attendance. The results show an increase in the mean from the pretest to the posttest.

Table 3. Descriptive statistics for SQ

	Mean	N	SD	Std. Error Mean
Q1	2.4894	126	.74092	.06601
Q2	3.2423	126	1.08039	.09625

The results of the paired sample t-test for the students' perceptions before and after their teachers' INSET classes (Table 3) signifies a statistically significant difference between the first and second answers to the questionnaire ($t(125) = -13.5$, $p < 0.01$). The mean increase in the scores is 0.75 in the posttest, with a 95% confidence interval ranging from -0.86 to -0.64. The eta squared statistic (0.06) indicates a moderate effect size. Therefore, the researchers can conclude that the students believed that the classes had affected their teachers' behavior positively. This result urges the researchers to verify the teachers' change of teaching performance after the classes (third research question).

Table 4. Paired sample T-test, students' perception of teachers' behavior before & after INSET courses

	Paired Differences				t	df	Sig. (2-tailed)	
	Mean	SD	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower				Upper
Q1 Q2	-.75283	.62241	.05545	-.86258	-.64309	-13.577	125	.000

Students' Pretest and Posttest

Table 5 shows the descriptive statistics of the students' performance on the language proficiency test before and after their teachers' INSET class attendance. As shown, the mean value of test 2 ($M=51.95$, $SD=7.35$) is larger than that of test 1 ($M=48.77$, $SD=7.30$), indicating the students exhibited better performance on the second test.

Table 5. Descriptive statistics, students' scores before & after teachers' INSET classes

	Mean	N	SD	Std. Error Mean
Pair 1 pre-test	48.77	126	7.302	.650
post-test	51.95	126	7.358	.656

The paired sample t-test for the students' test scores before and after their teachers' INSET classes (Table 6) reveals a statistically significant difference ($t(125)=-8.29, p \leq 0.01$). The mean difference is -3.183, with a 95% confidence interval ranging from -3.94 to -2.42. The eta squared statistic (0.35) shows a large effect size. That is to say, 35% of the knowledge students showed on the posttest is due to their teachers' change of behavior. Therefore, the answer to the fourth research question is positive, and the classes were beneficial for the teachers.

Table 6. Paired sample T-test between students' scores before and after the INSET class

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	SD	Std. Error Mean	95% Confidence Interval of the Difference				
				Lower	Upper			
Pair 1 pre-test post-test	-3.183	4.309	.384	-3.942	-2.423	-8.291	125	.000

Discussions

The study sought to evaluate the success of INSET classes in improving teachers' knowledge base. This issue is significant because of the paradigm shift in Iran that resulted in developing new high school textbooks. Since teachers need to go through developmental stages to construct their professional identity (Rashtchi & Khoshnevisan, 2019), paying heed to their training and examining the outcome of such training is crucial for teachers' success. The large effect size of the ELTKB questionnaire obtained from two administrations showed that grade 11 English language teachers' reaction (level one of Kirkpatrick's model) to the improvement of their knowledge base was constructive. That is to say, teacher participants found the content of the INSET classes useful in promoting the components (CK, TK, TCK, PK, PCK, TPK, TPACK) of their knowledge base. This finding is in line with that of Cahyono, Dwi Kournianti, and Mutiaraningrum (2016), Messina and Tabone (2012), Mouza and Wong (2009), Karimi (2011), Liu (2013), and Liu, Yu, Li, and Wen (2014) who reported the improvement of teachers' knowledge base after INSET classes.

The results of the interviews helped the researchers infer what teachers expected from the classes. Although they mostly believed that the classes had helped them learn about CLT and new ways of conducting classes, they did not find the classes adequate. They emphasized the shortcomings and provided ideas for the improvement of the classes. The researchers believe that authorities should consider the demands of modern life, the needs of the new generations, the development of technology, and thus modify their view regarding the INSET classes. The interviews showed that the classes are not in harmony with the needs of teachers and students. This view necessitates considering a more serious role for the use of new technology in the classes and training teachers to meet the needs of their students.

The results obtained from observations revealed grade 11 teachers' improvement in adopting new ways of teaching. The observations showed a change in student-teacher interactions, teachers' presentations, and students' practices in the classes. They also signified that the teachers were trying to put into practice what they had learned during the classes. They were more willing to establish rapport with their students; there was a drastic change in the assessment of the students, application of technology, and the use of English as the medium of teaching that would affect the students' learning. This finding is in line with Jansen (2014), who found that teachers used their learning experience,

teaching theories, and even students' conditions for delivering instruction after they understand what is expected from them. However, the researchers, through observations and in agreement with Bibi and Khan (2017), could not verify that the teachers could implement all components of the knowledge base in their classes.

The observations after the INSET classes indicated a promising start to infuse technology into language teaching in most of the grade 11 classrooms. Teachers' intention to use available technological instruments such as DVD players, tablets, and cell phones to teach English could be due to their participation in INSET classes. Moreover, during the observations, it was realized that teachers' avoidance of incorporating technology into teaching content was not for the lack of technological knowledge but due to the lack of technological devices. Besides, they did not receive sufficient support from school principals and administrators. The evaluation of teachers' success will be more accurate when they find the opportunity to employ what they have learned in their classes.

The results of the present study found support from Tseng, Cheng, and Lin (2011), who showed that teachers who received instruction could employ their knowledge in teaching grammar by using computer technology such as PowerPoint, Hot Potatoes, weblogs, and the Internet. Implementing new techniques and strategies in teaching grammar, vocabulary, and integrating technology into teaching practices implied that gaining new abilities had increased teachers' self-confidence. This finding is also reported by Powell, Terrell, Furey, and Scott-Evans (2003) and Davies and Preston (2002), who related teachers' quality of teaching to the improvement of their self-confidence due to the INSET class attendance.

Another dimension of this study was exploring the change of teachers' teaching behavior after attending INSET classes corresponding to the third level of Kirkpatrick's model. As Kiely (2009) states, the result of a program evaluation can be recognized through the investigation of knowledge or skills students achieve at the end of the program. The students' change of opinion before and after their teachers' attendance (signified by the large effect size) points to the efficacy of teacher training programs. The students' performance on the general English test (verified by large effect size), also, reveals that INSET classes not only could improve students' learning outcome (Hermans, Sloep, & Kreijns, 2017; Angrist & Lavy, 2001) but also could lead to productive results in teachers' performance.

Conclusion and Implications

The current four-level program evaluation study employed the Kirkpatrick's model (Kirkpatrick & Kirkpatrick, 2006) to examine the effectiveness of INSET programs in Guilan Province. The findings verified the usefulness of INSET programs in developing EFL teachers' knowledge base and fostering their professional growth. However, using recent advances in teaching methodology and technology use can make the classes more fruitful and can encourage teachers to look for innovative ways of instruction in their classes. Intensive training courses and workshops are excellent solutions for overcoming the problems teachers face for gaining self-efficacy. Also, authorities should focus on motivating teachers to integrate the use of technology into their teaching practice. However, training teachers should not be limited to official INSET programs. Virtual classes can provide the opportunity for interaction and exchange of information between teachers and instructors. They can promote teaching, assessment, lesson planning, classroom management, error correction procedures, and technology use. It is worth mentioning that an organized and pre-planned teacher training program that considers policymakers, INSET instructors, and teachers' interests into account is necessary. This study considered the perception of English language teachers. The researchers

suggest exploring the perceptions of INSET trainers, evaluation of the content of the INSET programs, and instructional materials in further studies. Additionally, teachers' characteristics such as age, university level, and teaching experience may provide more valuable information regarding teachers' perceptions of INSET classes.

Conflict of interest

The authors declare that there is no conflict of interest in this report. The study was carried out without funding from any institution.

Ethics

In gathering the data, in all stages of the study, ethical issues have been considered.

The data collected and used in this study are available and can be accessed via e-mail to maryammahmoudi75@yahoo.com.

References

- Angrist, J. D., & V. Lavy. (2001). Does teacher training affect pupil learning? Evidence from matched comparisons in Jerusalem public schools. *Journal of Labor Economics*, **19**(2), pp 343-369.
- Badu, S. Q. (2013). The implementation of Kirkpatrick's evaluation model in the learning of initial value and boundary condition problems. *International Journal of Learning & Development*, **3**(5), pp74-88. DOI:10.5296/ijld.v3i5.438
- Bates, R. (2004). A critical analysis of evaluation practice: The Kirkpatrick model and the principle of beneficence. *Evaluation and Program Planning* **27**, 341–347. DOI:10.1016/j.evalprogplan.2004.04.011
- Bates, R. & Coyne, T. H. (2005). Effective evaluation of training: Beyond the measurement of outcomes. *Human Resource Development Quarterly*, **11**(4), pp333-360.
- Best, J. W., & Kahn, J. (2006). *Research in education*. Chicago, IL: Pearson.
- Bibi, S., & Khan, S. H. (2017). TPACK in action: A study of a teacher educator's thoughts planning to use ICT. *Australian Journal of Teacher Education*, **33**(4), pp70-87.
- Birjandi, P., & Derakhshan Hesari, A. (2010). Teachers' perceptions of the present and optimum status of the in-service EFL teacher preparation programs. *English Language Teaching*, **3**(4), pp47-57. DOI: 10.5539/elt.v3n4p47
- Bledsoe, M. D. (1999). Correlations in Kirkpatrick's training evaluation model. (Doctoral Dissertation), University of Cincinnati, Accessed June 29, 2018, at <https://www.learntechlib.org/p/125181>
- Brewer, E. W. (2011). Evaluation models for evaluating educational programs. In V.C.X. Wang (Ed). *Assessing and evaluating adult learning in career and technical education*. pp.129-153. Hershey, PA: IGI Global.
- Cahyono, B.Y., Dewi Kurnianti, O., & Mutiaraningrum, I. (2016). Indonesian EFL teachers' application of TPACK in in-service education teaching practice. *International Journal of English Language Teaching*, **4**(5), pp16-30.
- Cohen, D., & Hill, H. (1997). *Policy, practice and learning*. Paper presented at the annual meeting of the American educational research association, March 1997, Chicago, IL.
- Davies, R., & Preston, M. (2002). An evaluation of the impact of continuing professional development on personal and professional lives. *Journal of In-service Education*, **28**(2), pp 231-254.
- Eghtesadi, A. R., & Hassanabadi, S. (2016). An evaluation of in-service training courses for the seventh grade English package for Iranian schools (Prospect 1). *Journal of Applied Linguistics and Language Research*, **3**(5), 130-143.
- Freeman, D. (2002). The hidden side of the work: Teacher knowledge and learning to teach. A perspective from North American educational research on teacher education in English language teaching. *Language Teaching*, **35**, pp1-13. DOI:10.1017/S0261444801001720
- Gill, M., & Sharma, G. (2013). Evaluation of vocational training program from the trainees' perspective: An empirical study. *Pacific Business Review International*, **6**(5), pp35-43.

- Grosso de Leon, A. (2001). *Higher education's challenge: New teacher education models for a new century*. New York, NY: Carnegie Corporation of New York.
- Hashemian, M., & Azadi, Gh. (2014). EFL teachers' understanding of the teaching portions of INSET programs, *RALS*, 5(1), 62-76.
- Hermans, F., Sloep, P., & Krteijns, K. (2017). Teacher professional development in the contexts of teaching English pronunciation. *International Journal of Educational Technology in Higher Education*, 14(23), 1-17. DOI: 10.1186/s41239-017-0059-9
- Jansen, A. (2014). Exploring non-native EFL teachers' knowledge base: Practices and perceptions. *International Journal of Applied Linguistics & English Literature*, 3(6), pp252-259.
- Karimi, M. N. (2011). Variations in EFL teachers' pedagogical knowledge base as a function of their teaching license status. *The Journal of Teaching Language Skills*, 3(3), pp83-114.
- Kazemi, A., & Ashrafi, M. (2014). In-service training programs for Iranian EFL teachers revisited. *International Journal of Asian Social Science*, 4(10), pp1062-1076.
- Khanjani, A., Vahdany, F., Jafarigohar, M. (2016). EFL teacher education in Iran: Does it promote trainees' pedagogical content knowledge? *Journal of Research in Applied Linguistics*, 8(2), pp159-186.
- Kiely, R. (2009). Small answers to the big question: Learning from language program evaluation. *Language Teaching Research*, 13(1), 99-116. DOI: 10.1177/1362168808095525.
- Kiely, R., & Rea-Dickins, P. (2009). Evaluation and learning in language programs. In K. Knapp, B. Seidlh, & H. G. Widdowson (Eds.). *Handbook of language communication and learning* (pp. 663-694). Berlin: Walter de Gruyter.
- Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels* (2nd ed.). San Francisco: Berrett-Koehler.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs: The four levels* (3rd ed.). San Francisco, CA: Berrett-Koehler.
- Liu, S. (2013). Pedagogical content knowledge: A case study of ESL teacher educator. *English Language Teaching*, 6(7), pp128-138.
- Liu, S., Liu, H., Yu, Y., Li, Y., & Wen, T. (2014). TPACK: A new dimension to EFL teachers' PCK. *Journal of Education and Human Development*, 3(2), pp681- 693.
- Messina, L., & Tabone, S. (2012). Integrating technology into instructional practices focusing on teacher knowledge. *Social and Behavioral Sciences*, 46, pp1015-1027.
- Mouza, C., & Wong, W. (2009). Studying classroom practice: Case development for professional learning in technology integration. *Journal of Technology and Teacher Education*, 21(5), pp509-523.
- Powell, E., Terrell, I., Furey, S., & Scott-Evans, A. (2003). Teachers' perceptions of the impact of CPD: An institutional case study. *Journal of In-Service Education*, 29(3), 389-404.
- Roberts, J (1998). *Language teacher education*. London: Routledge.
- Rashtchi, M., & Khoshnevisan, B. (2019). The developmental stages of teachers: A critical analysis. In B. J. Wayne & C. Cobanoglu (Eds.). *Advances in global education and research* Vol. 3, Sarasota, Florida: ANAHEI. pp. 2-8
- Spitzer, D., & Conway, M. (2002). *Link training to your bottom-line*. Infoline. Alexandria, VA: ASTD.
- Strauss, A., & Corbin, J. (1990). *Basics of quantitative research: Grounded theory procedures and techniques*. Newbury Park, CA: Sage.
- Swanson, R. A. (2001). *Assessing the financial benefits of human resource development*. Cambridge, MA: Perseus.
- Tseng, J., Cheng, Y., & Lin, C. (2011). *Unraveling in-service EFL teachers' technological pedagogical content knowledge*. *The Journal of ASIA TEFL*, 8(2), pp45-72.
- Weiss, C. H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), pp21-33. DOI: 10.1177/109821409801900103
- Wertz, C. (2005). *Evaluation of CLAD training in northern California*. (Doctoral dissertation). University of Southern California, Accessed June 29, 2018, from ProQuest Dissertations Publishing,
- Yusoff, M. A. M., Ahmad, J., Mansor, A. N., Johari, R., Othman, K., & Che Hassan, N. (2016). Evaluation of school based assessment teacher training program. *Creative Education*, 7, pp 627-638. DOI: 10.4236/ce.2016.74065

APPENDICES

Appendix A: English Language Teachers' Knowledge Base Questionnaire (ELTKB)

The purpose of this questionnaire is to examine the effect of INSET courses on English teachers' knowledge base. Please answer the questions, sincerely. All responses will be kept confidential. For any queries, do not hesitate to contact maryammahmoudi75@yahoo.com. Thank you for your time and patience.

Section I. Bio-data: In this section, please choose the option that applies to you.

1. Gender

1. Male 2. Female

2. Teaching experience

1. Less than 5 2. Between 5 to 20
3. Between 10 to 15 4. More than 15

3. The school grade you teach

1. Junior high school 2. Secondary high school

4. University degree

1. BA 2. Student of MA or MA
3. Ph.D. candidate or Ph.D.

5. The number of INSET courses you have participated since 2012

1. Two to three times 2. Three to five times
3. More than five times

Section II: The following questionnaire examines the effect of INSET programs on English language teachers' knowledge base (Content Knowledge, Pedagogical Knowledge, Pedagogical Content Knowledge, Technological Knowledge, Technological Content Knowledge, Technological Pedagogical Knowledge, Technological Pedagogical and Content Knowledge). The questionnaire is a 5-point Likert-type scale, 1 (strongly disagree = SD), 2 (disagree = DA), 3 (neutral = N), 4 (agree = A) and 5 (strongly agree = SA). Please read the items carefully, and then tick the option that best represents your position.

INSET classes helped me ...	5 SD	4 D	3 N	2 A	1 SA
1. improve my English grammar.					
2. improve my pronunciation.					
3. improve my reading comprehension.					
4. improve my listening skills.					
5. improve my writing skills.					
6. use a computer, projector, and electronic board.					
7. use office programs such as Word, PowerPoint, and the like.					
8. solve primary problems that may happen to electronic devices such as a computer, printer, scanner, and electronic board.					
9. attach video, picture, or text.					
10. use electronic devices such as a computer, laptop, cell phone to increase my reading skills.					
11. use electronic devices to increase my listening skills.					
12. use electronic devices to learn pronunciation and falling and rising of sentences.					
13. use grammar training software to increase my grammatical competence.					

INSET classes helped me ...	5 SD	4 D	3 N	2 A	1 SA
<p>14. understand the cultural differences between my native language and second language with the help of watching a movie.</p> <p>15. improve my formal and informal writing skills with the help of electronic devices and training software.</p> <p>16. improve my vocabulary knowledge with the help of training software.</p> <p>17. know the different learning styles of my students.</p> <p>18. involve shy and introvert students in the classroom discussions.</p> <p>19. manage students' group work.</p> <p>20. explain the lesson according to the needs of the students.</p> <p>21. do reflection after my teaching.</p> <p>22. use a variety of techniques for teaching language skills (listening, speaking, reading, writing).</p> <p>23. use different teaching materials (flashcards, pictures, graphs, etc.).</p> <p>24. consider students' differences in their level of intelligence.</p> <p>25. communicate with the school principal and parents to enhance my students' learning.</p> <p>26. use the experience of other teachers in teaching.</p> <p>27. use different techniques (oral questions, written questions, role play, class activity, etc.) to assess their learning.</p> <p>28. provide a stress-free environment for my students.</p> <p>29. teach vocabulary with its phonology.</p> <p>30. assess the accuracy of my students' pronunciation.</p> <p>31. teach grammar explicitly and implicitly.</p> <p>32. evaluate my students' grammatical knowledge in each lesson.</p> <p>33. consider vocabulary instruction in teaching reading skills.</p> <p>34. manage time for teaching each skill (listening, speaking, reading, and writing).</p> <p>35. compare native and foreign language conversations in teaching the listening skill to make them more understandable for my students.</p> <p>36. perform before, while, and after activities for teaching each skill properly.</p> <p>37. recognize students' problems after teaching each skill and solve them.</p> <p>38. use different techniques and strategies to involve all students while teaching each skill (listening, speaking, reading, writing).</p> <p>39. evaluate students' ability after teaching each skill (listening, speaking, reading, writing).</p> <p>40. evaluate my teaching after teaching each skill (listening, speaking, reading, writing).</p> <p>41. recognize my students' learning style (audible, visual, kinesthetic, etc.) with the help of electronic devices (DVD player, CD player, computer, etc.).</p> <p>42. involve students in out of class activities and projects to increase their learning opportunities.</p> <p>43. use various electronic devices to teach each skill differently.</p>					

INSET classes helped me ...	5 SD	4 D	3 N	2 A	1 SA
44. increase students' learning to communicate with the principal and their parents with the help of e-mail, telegram, and text message.					
45. assess the students' learning with the help of electronic devices.					
46. reduce the anxiety level of the students with the help of electronic devices.					
47. teach pronunciation of the words with the help of electronic devices (computer, laptop, CD player).					
48. teach falling and rising of the sentences with the help of electronic devices (computer, laptop, CD player).					
49. assess the students' grammar knowledge with the help of electronic devices.					
50. use teaching grammar with the help of electronic devices					
51. use pictures, graphs, and voice to brainstorm students in teaching reading skills.					
52. use electronic devices to teach vocabulary					
53. use electronic devices to perform before, while and after activities of the teaching of each skill					
54. assess my students' vocabulary knowledge with the help of electronic devices					
55. evaluate my teaching of vocabulary, grammar, and all skills (listening, speaking, reading, and writing) through recording of my classes.					
56. enhance my students' writing competence by applying different techniques of teaching with the help of electronic devices.					

Appendix B: Interview Questions

1. Could the INSET course help you improve your language skills (reading, speaking, listening, and writing)?
2. Could the INSET course help you improve your technological knowledge (using a computer, laptop...) in your teaching?
3. Could the course help you find a better way for
 - a. managing your classes?
 - b. using different techniques and methodologies for teaching language skills?
 - c. using different types of assessing your students' learning?
4. Could the course help you mix technology in teaching language skills?
5. Could the course help you do reflection after teaching?
6. Have you encountered any problems implementing what you have learned in these classes?
7. Have the INSET classes had any NEGATIVE impact on your classroom, your students, or you professionally?
8. Do you have any other ideas or suggestions that you would like to share?

Appendix C: Observation Checklist

Yes To some extent No

-
1. The English language is the medium of instruction.
 2. Grammar is taught implicitly.
 3. Brainstorming is used to teach new words.
 4. A speaking dictionary is used to check the pronunciations.
 5. Technologies are used to teach listening.
 6. Group work activities are performed.
 7. The class is learner-centered.
 8. Meaningful language learning is the focus of instruction.
 9. Four skills are integrated.
 10. Students' participation is welcomed.
 11. Students' comprehension is checked during teaching.
 12. Sufficient time is allowed to complete tasks and group work activities.
 13. Conflicts or different ideas are mediated.
 14. Different learning styles are considered.
 15. Periodic feedback is applied.
 16. Teachers' overall behavior is encouraging.
 17. Educational technology is used appropriately.
-

