The effect of in-service English education on medical professionals’ language proficiency

Samad Sajjadi¹, Majid Ahmadi², Maryam Heidarpour³, Ali Salahi Yakta⁴, Naghmeh Khadembashi⁵, Mohammad Rafatbakhsh⁶

¹Associate Professor, Department of Applied Linguistics at the English Language, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ²Associate Professor, Department of Applied Linguistics at the English Language, School of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ³Instructor, Department of Medical and General English Courses at the English Language, School of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ⁴Instructor of Medical and General English Courses at the English Language Department, College of Paramedical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran. ⁵Instructor, Department of Medical and General English Courses at the English, Paramedical College, Shiraz University of Medical Sciences, Shiraz, Iran.

Background: Despite its inevitable significance, the effect of in-service English education on medical professionals has rarely been studied longitudinally. The reason can be issues such as physicians’ heavy workload, commuting problems, inappropriate class times, and inexperienced teaching staff. Materials and Methods: A needs assessment worksheet was administered to faculty members of Shahid Beheshti University of Medical Sciences in Tehran and the responses were analyzed. A project for the promotion of faculty members’ English proficiency was formulated. Then, following a placement test, 235 applicants from the university colleges and hospitals were classified into 28 homogeneous groups. After four terms of instruction, the participants’ scores on the pre- and post-assessments were analyzed. Results: There was significant improvement in participants’ total scores on different communicative skills (P<0.001). Regarding individual skills also, they achieved meaningful gains on listening (P<0.001), writing (P = 0.038), and grammar (P<0.001), but failed to progress significantly on reading comprehension (P = 0.523). Conclusion: The administration of in-service education for skill-oriented courses, over a long period, can be quite encouraging and should be further strengthened. Regular instructions on each individual skill on the one hand and on their combination on the other are essential for success in such education.

Key words: English language skills, in-service education, medical professionals

INTRODUCTION

In-service education facilitates the enhancement, development, and updating of knowledge and professional skills.[1] It offers the possibility of obtaining new skills and qualifications or specialization, professional training or vocational experience acquired earlier. It additionally provides a unique opportunity for individual self-study, irrespective of one’s age.[1,2] Education has a significant effect on blood pressure self-monitoring in patients with hypertension.[3]

Numerous studies have investigated the way in-service education functions in different professional settings such as teaching hospitals and colleges. In-service education program can result in short-term behavioral changes in educators’ use of abstract language and print references.[1] It gives the instructors the necessary knowledge of language concepts related to early literacy instruction, which they could then integrate into their pre-service reading courses.[4] The outcome of in-service training on language facilitation strategies of childcare providers in day care centers has been referred to in Girolametto’s study[7] in which sixteen caregivers were randomly assigned to experimental and control groups. Caregivers were taught to be responsive to children’s initiations, engage children in interactions, model simplified language, and encourage peer interactions. At post-test, the experimental group waited for children to initiate words, engaged them in turn-taking, used face-to-face interaction, and included unwilling children more frequently than the control group. In turn, children in the experimental group talked more, produced more combinations, and talked to peers more often than the control group.[7]

McKeon’s work provided new insights into the ways by which professional pedagogical learning has developed.[8] Shawer’s research examined the effect of classroom teacher development program on EFL (English as-a-foreign language) college teachers and students. According to the results, the program significantly improved student-teacher subject pedagogical and curricular skill, resulting in professional satisfaction.[9,10] Sabouri et al.[11] indicated the need for the development of new instructional materials with properly ordered visuals as part of their ingredients for educational settings, such as the in-service education. The use of new materials with their accompanying pictures in such settings, according to the authors, could help to achieve optimal comprehension.
The studies cited above are a few samples of numerous pieces of research carried out thus far to investigate the role of in-service education in enhancement, development, and updating of knowledge and professional skills from different perspectives. Although these studies form a vast array of work, there are still many questions waiting for answer. A significant issue requiring further investigation, for example, is the role of such instruction in language skills of medical professionals visiting patients in EFL settings. Physicians’ mastery over such skills is of special significance, as the skills can serve as communication means by which physicians and patients relate to each other to achieve therapeutic goals.\(^\text{[12]}\)

In line with such issues, the current study specifically investigated the impact of classroom-level language instructions on medical professionals’ skills development. It intended to know how successful in-service instructions could be for medical professionals, serving as faculty members at Shahid Beheshti University of Medical Sciences. The study addressed the following questions:

- Would it be practically feasible to carry out a large-scale in-service English education to enhance the language proficiency of medical professionals working at numerous hospitals and medical colleges located wide apart?
- Will a regular in-service English education elevate physicians and faculty members’ command of English on each individual skill of writing, listening, grammar, and reading in the same way?

Based on the research questions, the following hypotheses are formulated: (a) In-service English education can be successfully administered to develop the language proficiency of medical professionals working at numerous hospitals and medical colleges located wide apart. (b) In-service English education will develop faculty members’ command of English on each individual skill of writing, listening, grammar, and reading in the same way.

**MATERIALS AND METHODS**

**Participants**
Participants were 235 volunteer faculty members from different hospitals and colleges of Shahid Beheshti University of Medical Sciences in Tehran. Of this total, 99 volunteers (42%) were females and 136 volunteers (58%) were males. They participated in a placement test a few days before the commencement of the program. Based on the test results, they were divided into 28 relatively homogeneous groups, each with 8 to 14 members. The groups were composed of four different proficiency levels: pre-intermediate, intermediate, upper-intermediate, and advanced. The reason for dividing participants into different groups based on their proficiency level was to have linguistically homogeneous participants together in similar classes, which could help achieve more successful instructions. The participants belonged to different medical orientations and specialties; nonetheless, the main criterion for their classification into particular learning groups was their command of English, measured by a placement test carried out before starting the classes.

**Venue of instructions**
The classes were held in the faculty members’ work places in different colleges and teaching hospitals. This could help save a lot of time, by avoiding commuting problems in a crowded city like Tehran. The number of classes in each place varied from 1 to 4, depending on the number of applicants in the workplace. If the level of the class in an applicant’s workplace did not match his proficiency level, he could attend an appropriate class in another hospital or college at the closest vicinity. Performing the instructions in the applicants’ workplace made the training very cost beneficial; the faculty members did not have to pay high tuitions for tutorial staff. Nor did they have to pay for commuting charges; they saved a lot of time by attending the training sessions at the closest distance in their affiliate colleges or hospitals.

**Instructors**
The teaching staff, undertaking the instructions for four successive terms, were mostly recruited from a language institute in Tehran, based on a contract agreed upon and endorsed by the English Language Department of the university and the head of the Institution. The instructors’ qualifications had to be endorsed by a team of experts from the English Language Department of the university before they could start instructing at one of the teaching hospitals or colleges. The reason for recruiting the instructors from an institution rather than the English Language Department was staff limitation in the Department. The selection of the institution and the instructors followed a detailed process. The main criteria were (a) the instructors’ teaching standards and qualifications, measured by class observation, and (b) the fees they demanded for each teaching session.

**Instruments**
The instruments used in this study were as follows: (a) A Needs Assessment Worksheet; (b) an objective placement test, versions “A” and “B”; and (c) a Course Information Sheet.

*The Needs Assessment Worksheet*
The Worksheet, with 20 items, was developed based on Munby’s “needs analysis” model.\(^\text{[13]}\) Although Munby’s model developed during the 1970s, “it has been [according
to Jordan, p.24] very influential: either developments have stemmed from it, or as a result of reactions to it.\[14\]

The Needs Assessment Worksheet was not required to be subjected to variability measures. The reason is that it was not intended to investigate a latent variable as the items included in it did not represent a “shared” variance. They varied in the number of options they had—from two to five—depending on the type of item.

The Needs Assessment worksheet was filled out by a total of 450 faculty members from all colleges and teaching hospitals before the commencement of the experiment. The items forming the worksheet dealt with issues such as the significance of English in faculty members’ career, the type of skills they were more interested in, the role of standard examinations such as International English language testing system in the future of their work, their communication purposes, and their desired level of proficiency.

**Objective placement test**

The objective placement tests, in two parallel versions (“A” and “B”), each with 70 multiple-choice items, had been designed by Lesley et al.\[15\] The tests followed two main objectives of (a) classifying the participants into homogeneous groups, with similar levels of English proficiency, and (b) serving as pre-test/post-test instruments (in two parallel forms), to identify score gains of the subjects under instruction. They measured receptive skills of listening, reading, and grammar recognition. Also, they included composition topics to assess participants’ writing skills. All parts of the tests were based on the objectives, content, and language of different levels of Interchange Third Edition and Passages.\[16\] In section 1, the Listening section with 20 items, the participants listened to nine conversations and answered one or more questions about each one. They heard each conversation only once. Section 2, the Reading section, had several short passages, with 20 test items. The participants were expected to choose the correct answer for each item based on the information in the accompanying passage. In section 3, the Language Use section with 30 items, the participants chose the correct completion for each item and marked their choice on the answer sheet. They were allowed a total of 50 minutes to complete the test, with 15, 20, and 15 minutes for each of the Listening, Reading, and Language Use sections, respectively. In its Essay Form, the test wants the test takers to write an organized essay about one of the topics presented in this section.

The reasons for the selection of Lesley’s Placement test were that (a) the institution, where the teaching staff came from, used the Interchange Third Edition as their source materials to teach and practice English; (b) the same materials, i.e. the Interchange series, served as teaching materials for the research project as well; and (c) the test was recommended by EFL teachers and test developers who were familiar with the research objectives of the project and its subject selection procedure.

**The Course Information Sheet**

The Course Information Sheet was intended to find out whether all the post-test participants had already taken the pre-test; whether they attended the classes regularly for four successive terms; and whether they received any skills-oriented English instructions other than those arranged by the Project. Depending on the participants’ answers to these questions, due measures were adopted to avoid cases that could undermine the quality of the study. These measures were adopted before ending the terms.

**Procedure**

The participants attended the pre-test examination a few weeks before the start of the classes. The place where the tests were carried out was quite light and spacious, with very appropriate acoustics. Based on the results of the test, the subjects were classified into 28 smaller groups, each with 8 to 14 participants. Variations in the number of participants attending each class were due to their proficiency level and their vicinity to the venue where the classes were held. Each term, having a total of 25 sessions, with two sessions per week, took nearly three months to conclude. During the term, as part of the terms and conditions indicated in the contract endorsed by the institution supplying instructors and the English Department, all the classes were observed by faculty members from the English Department to make sure of high quality of instructions. Course assessment included midterm and final examinations and class activities during the term.

The examination committee did their best to have the post-test in the same venue where the pre-test was already held. Nonetheless, due to a number of problems, particularly participants’ difficulty in agreeing upon a common time to assemble, the examination committee decided to hold the post-test in three venues simultaneously, instead of one. However, due measures were adopted to make the post-test venues similar to that of the pre-test (in terms of acoustics, space, light, type of chairs, etc). The selection of venues for the post-test was subject to the research committee’s approval. The instructors were also asked to (a) justify the class participants of the importance of their attendance in the post-test and (b) make due arrangements for their presence in one of the three examination venues. In terms of the examination questions also, as already indicated, there was full conformity between the two examinations as the post-test was indeed a parallel version of the pretest.\[15\]
All the training was conducted early in the morning, before the start of work, or in the afternoon, right after the office hours. In few cases, some adjustments were made to the participants’ pre-established schedules to find appropriate times for the classes.

The level of the classes and their frequency per week depended on the physicians’ level of English proficiency and the urgency they faced in acquiring the skills. There were two sessions per week for different proficiency levels. Nonetheless, if the participants themselves, or their affiliate colleges or hospitals, wanted to have more sessions per week, due measures were adopted to arrange more sessions for them. In one case, for example, some volunteers had little time to get ready for the IELTS examination, which was essential for passing their fellowship abroad; the research committee, accordingly, arranged more than two sessions per week for them.

Following the administration of the post-test, three members of the Research Committee undertook the correction of examination papers. The multiple choice parts of the examination, including reading, grammar, and listening were easy to correct but the writing part was time taking and difficult as it lacked any pre-specified key. Hence, to avoid correction bias, the Research Committee agreed upon a set of correction criteria. Right before the correction procedure, the writing papers were photocopied in order to have sufficient copies for each of the three members of the correction team. To avoid correction biases, the participants’ names were replaced by numerical codes to keep them anonymous. Following the correction procedure, the three scores assigned to each composition by the correction team were added up and averaged to obtain a single score for each participant’s writing. The challenging issues in each composition were discussed and agreed upon by the committee members before finalizing each participant’s score. The correction criteria were identical for both the pre-test and post-test compositions.

With the conclusion of correction procedure, there were five groups, or 10 sets, of scores to be compared, as follows: (a) pre- and post-test grammar scores, (b) pre- and post-test listening scores, (c) pre- and post-test reading comprehension scores, (d) pre- and post-test composition scores, and (e) pre- and post-test totals. The scores were then subjected to a number of statistical computations aimed at comparing participants’ scores on the pre-test and post-test.

Statistical Analysis
A number of statistical computations, using SPSS 16.0 software, were carried out to analyze the data as follows: (a) standardization of scores by converting them to new figures, i.e., of 20, (b) application of one-sample Kolmogorov-Smirnov test for assessing the normality assumption of variables and (c) paired t-tests for significant contrasts. Paired t-tests were intended to compare the pre- and post-assessment scores for the skills under study. P value <0.05 was considered statistically significant.

RESULTS
A total of 264 faculty members participated in the pretest, of whom 235 members continued the course to the end. Altogether 11 subjects (4.47%) did not finish the course, which is well below the allowed figure of 10% for dropouts. Upon the completion of the project, 40 subjects were randomly asked to sit for the post-test examination. A post-test, with fewer randomly selected subjects, allowed the researchers to save time and money in preparing the test booklets and administering the post-test. This way the researchers would also find it much easier to set a convenient date for the post-test as they had to make fewer arrangements with those who were to attend the post-test examination.

Following the post-test assessment, the examination papers were corrected and the data were sorted out and tabulated for further computations. According to the Kolmogorov-Smirnov test, the P value for variables of writing, listening, grammar, and reading comprehension were all within the pre-specified statistical limits (i.e., P>0.05), justifying the application of parametric tests. Then, a set of paired t-tests were used to compare the mean score differences for the skills on the pretest and post-test examinations. According to the results, four of five mean scores (rows 1,2,3,5, Table 1) were significantly different from one another. The mean score differences for the reading comprehension (row 4, Table 1), however, failed to be statistically significant, although the difference indicated some gains (P = 0.523).

The computations went further to measure the effect size too. [Table 1] also reports the effect sizes whose base of interpretation is Dörnyei,10 that is, eta squared is interpreted as 0.01 = small effect, 0.06 = moderate effect, 0.14 = large effect. The mean score increments and effect size magnitudes were noticeable for different skills. The mean writing skill score (9.80) increased significantly to 10.45 after training (P = 0.038) with a moderate magnitude of 0.06 for effect size. Listening comprehension skill also increased meaningfully from 9.02 to 11.52 (P<0.001) with a large magnitude of 0.18 for effect size. The mean grammar score (16.22) increased significantly to 19.20 with a very large magnitude of 0.25 for effect size. The increase in mean reading comprehension (from 14.17 to 14.55 with effect size of 0.01), however, was not statistically significant (P = 0.523).
The total score (49.21), representing participants overall scores on all the four skills together, increased to 55.72 (P<0.001) with a very large magnitude of 0.21 for effect size.

**DISCUSSION**

The participants had score gains on all the four skills under study and the gains on three skills were statistically significant. The effect sizes of 0.18, 0.25, and 0.06 for listening comprehension, grammar, and writing indicated the strength of the difference between pretest and post-test scores. The effect size of 0.21 for the total score also indicated the effective role of in-service instructions on subjects' overall proficiency in English. The reading comprehension with the effect size of 0.01, nonetheless, did not manifest a noticeable progress.

The results confirm the first hypothesis on effectiveness of the training which improved faculty members' communication skills meaningfully, without disrupting their normal flow of work. The progress made following the education is in line with previous findings indicating that on-the-job training could transfer skills and knowledge.[19] The progress also supports the findings in Ancel's experiment[10] according to which nurses' scores increased from 155.6 to 180.5, following an in-service training. Nurses' empathic skills also elevated with regard to all variables under instruction.[10]

The training, as the results indicate, can be used to successfully target each skill individually (such as the listening or writing) or a group of skills together (such as the total scores). Previous studies also successfully addressed the lexicon items, as an individual skill, essential to the profession.[19] Targeting an individual skill could help healthcare professionals improve their communication skills gradually without being subjected to a heavy workload.[20]

A relatively poor record on the reading comprehension, in comparison with other skills, was at odd with the second hypothesis which predicted similar progress for each individual skill of writing, listening, grammar, and reading. This was apparently due to the insufficiency of instructions allocated to reading comprehension. To check the truth of this speculation, the research committee asked for the subjects and instructors' feedbacks on poor record of reading on the one hand and examined the textbooks and teaching materials on the other. They realized that the reading skill had received less attention than necessary, mainly because the class activities were more conversationally oriented; the textbooks, likewise, did not offer sufficient reading comprehension practice, as they were also more conversationally inclined. Hence, if a given skill is to develop sufficiently, it has to be given due attention, by receiving sufficient level of practice. One way of achieving this is to maintain a proper balance in the number and type of exercises, when writing a syllabus for a particular course, which could help to get the desired outcome for all the skills under instruction.

Lack of significant progress in a skill, such as the reading comprehension, has already been indicated in previous research.[6,18] In Oz's work,[6] for example, training brought about meaningful progress on emphatic communication skills but did not cause a noticeable change in emphatic tendency levels.[6] In Sullivan et al.'s study, in-service training enhanced subjects' knowledge noticeably, but did not meaningfully change their skills.[18]

The achievements of this study challenges the claim that in-service training often updates knowledge rather than skills.[18] The reason, according to the proponents of the claim, is the difficulties involved in providing suitable training conditions for skill-oriented courses. Nonetheless, significant progress in the skills under training in the current study rejects such claims. The reason for effectiveness of the training could be the provision of a series of measures as follows: needs assessment, placement test, class observation, classes at workplace, regular assessment and feedback, well-trained instructors, and proper teaching materials.

Needs assessment was quite essential to the success of the training. Before formulating the policies, the researchers performed a needs assessment survey in order to understand the situation and define the needs. They then used the results of the assessment to develop the training program. The program started with a placement test which was designed to measure participants' entry-

---

**Table 1: Comparison of pre-test and post-test skill scores, using paired t-test**

<table>
<thead>
<tr>
<th>Skills</th>
<th>Average out of 20</th>
<th>Difference</th>
<th>T</th>
<th>P value</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-test</td>
<td>Post-test</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing</td>
<td>9.80</td>
<td>10.45</td>
<td>0.65</td>
<td>1.76</td>
<td>-2.160</td>
</tr>
<tr>
<td>Listening</td>
<td>9.02</td>
<td>11.52</td>
<td>2.50</td>
<td>3.76</td>
<td>-4.203</td>
</tr>
<tr>
<td>Grammar</td>
<td>16.22</td>
<td>19.20</td>
<td>2.98</td>
<td>3.46</td>
<td>-5.219</td>
</tr>
<tr>
<td>Reading</td>
<td>14.17</td>
<td>14.55</td>
<td>0.38</td>
<td>3.67</td>
<td>-0.645</td>
</tr>
<tr>
<td>Total</td>
<td>49.21</td>
<td>55.72</td>
<td>6.51</td>
<td>8.65</td>
<td>-4.635</td>
</tr>
</tbody>
</table>
level skills in writing, reading, listening, grammar, and vocabulary. Placement scores served as a basis for placing the participants with relatively similar command of English in appropriate in skill-based courses. This way the participants attending the classes possessed similar abilities of English as an essential requirement for the success of skill-based courses.

Regarding class observation, a team of EFL experts from the English Department of the university regularly observed the instructors’ performance and provided objective behavioral feedback about it. The feedback was based on observed, factual, specific work-related behaviors, and working relationships. The feedback helped the instructors sustain good performance, develop new strategies, and improve performance when necessary.

The venue of the classes was carefully chosen. All the classes were held at workplace of faculty members before and after normal work hours. The instructors referred to the affiliate colleges or hospitals and led the instruction for a period of 100 minutes per session. Workplace classes made the training very cost beneficial for both the university and the faculty members who were able to schedule their training at the time most convenient to them. They did not have to commute long distances to attend the classes.

Assessment and feedback carried out regularly played an effective role in achieving ends. They went hand in hand with clear course aims and objectives. The feedback provided following each particular assessment proved to be very effective in facilitating a dialogue between lecturers and class participants. This was in line with Biggs’ remarks on the facilitative role of regular assessment and feedback.[3]

The teaching staff played a very significant role in meeting the program’s objectives; the teachers undertaking the instructions were selected from the teaching staff of a language institute following class observation; this made it possible for the research committee to recruit the most competent teachers who helped to improve the quality of educational processes. The teaching materials, Interchange series,[16] were also quite encouraging. They effectively combined topics, functions, and grammar; they covered the four skills of listening, speaking, reading, and writing, in addition to vocabulary and pronunciation. The self assessment Progress check let the participants check their skills in need.

Despite all its merits, this research, like many other studies, had its own limitations. A major issue was the correction of compositions which was difficult to do; due measures were adopted to tackle this issue by identifying a set of criteria based on testing principles for correcting “texts” and “compositions”; each composition was corrected and scored by three raters; the scores were then added up and averaged to obtain the final score of the composition. Another challenge in such studies is the possibility of numerous dropouts, which fortunately did not happen in the current study, owing to the importance of such classes to the participants and the measures adopted by the course organizers, as indicated above, to make class presence as convenient as possible. Among such measures, discussed above, the commuting issue was of special importance for the participants which was sorted out successfully by holding the classes at the participants’ workplace at an appropriate time of the day agreed upon by the participants themselves.

**CONCLUSION**

In-service education can significantly promote faculty members’ overall English proficiency; likewise, it can enhance their mastery over each particular skill, such as listening, writing, and grammar, in an EFL environment. Hence, given the increasing importance of English as an international means of communication, in-service education should be given due attention, by providing regular training and retraining of the professionals and staff members. This can guarantee our professionals’ appropriate mastery over the skills in need.

**ACKNOWLEDGMENTS**

The authors would like to thank Dr. Seyed Jalil Hosseini, the Vice Chancellor for educational affairs during the administration of the study, for his valuable insights and administrative supports. We also thank Dr. Mohammad-Reza Sohrabi for his constructive views on formulating and performing the study. Special thanks go to the university’s offices of Education and Research and the Paramedical College’s Research body for approving the research and, accordingly, providing the necessary funds for its execution. We should thank all the teaching staff who energetically carried out the instructions with special interest and dexterity. Special thanks go to all the medical professionals and faculty members from the university colleges and teaching hospitals who eagerly attended the classes, thereby making the execution of the research project possible.

**REFERENCES**


Source of Support: University’s offices of Education and Research and the Paramedical College’s. Conflict of Interest: None declared.