



## Evaluation of Proficiency Test Preparation Course Program at the Language Laboratory of FKIP-UGN

Yulia Rizki Ramadhani<sup>1,4</sup>, Ambiyar<sup>2</sup>, Ishak Aziz<sup>3</sup>

<sup>1</sup>Doctoral Program of Educational Sciences, Universitas Negeri Padang

Jalan Prof Dr Hamka Kampus Air Tawar, Padang 25131, Indonesia

<sup>2</sup>Faculty of Engineering, Universitas Negeri Padang

Jalan Prof Dr Hamka Kampus Air Tawar, Padang 25131, Indonesia

<sup>3</sup>Faculty of Sports Science Universitas Negeri Padang

Jalan Prof Dr Hamka Kampus Air Tawar, Padang 25131, Indonesia

<sup>4</sup>Department of English Language Education, Universitas Graha Nusantara

Kampus I : Bukit Tor Simarsayang, Padangsidempuan 22717, Indonesia

Corresponding email: [yuliadamanik44@gmail.com](mailto:yuliadamanik44@gmail.com)

---

### ABSTRACT

This study uses the Kirkpatrick evaluation model to evaluate the implementation of the Proficiency Test Preparation Course Program at the UGN FKIP Language Laboratory. This research is qualitative and quantitative descriptive research, with the type of research used evaluative research. The subjects of this study were 30 students who attended language training from various study programs at Universitas Graha Nusantara. The instrument used is a questionnaire with a Likert scale, a formal test through pre-test and post-test. Data analysis was carried out by drafting, classification, processing, interpretation and conclusion. The results show that the Proficiency Test Preparation Course Program, run at the Language Laboratory of FKIP UGN, can be continued because it helps non-English Department students develop foreign language competence. The results of this evaluation can be used as material for consideration in determining steps that can be taken to improve the implementation of the Proficiency Test Preparation Course Program so that the results of this program can be achieved optimally.

**Keywords: evaluation, proficiency test, Kirkpatrick model**



This work is licensed under Creative Commons Attribution License 4.0 CC-BY International license  
ISSN 2655-9722, DOI: 10.30650/ajte.v5i2.3638

---

# 1. INTRODUCTION

## 1.1. Introduction

National education in Indonesia has made various efforts to face the era of global competition in producing graduates who can compete and take advantage of opportunities to continue their education to higher education or the world of work. The efforts made by the government through the national education ministry are also supported by independent efforts made by schools and tertiary institutions to improve the quality of education and form a generation that has broad insights to compete in the era of globalization. In order to be able to compete in the age of globalization, it is necessary to reorient an innovative, adaptive, and collaborative curriculum, develop teaching materials, facilities, and administration and improve human resources that will carry out and support facing global competition.

One of the challenges in increasingly fierce global competition is how to increase the nation's competitiveness in increasing quality work and being able to compete so that the nation's progress can be achieved. This progress can be realized with a quality learning process that produces broad-minded, professional, superior, and far-sighted (visionary) graduates with the confidence and high self-esteem. Among the low competitiveness of human resources and the responsibility of universities is the problem of mastering foreign languages, especially English. Foreign language skills also contribute to improving the quality of educational graduates. The ability to speak foreign languages is something that college graduates must possess to compete in facing the globalization of the world of work (Rofaida & Gautama, 2019; Thariq et al., 2021; Ucu, 2021). Foreign language skills are needed to support the qualifications of graduates ready to enter the globalized world of work.

Mastery of English is crucial as an effort to increase competitiveness, both in the world of work and in the academic world. In the world of work, mastery of a foreign language is an added point and even becomes the main requirement for being accepted to work in a company, especially abroad. In the academic world, many still do not enter college or do not pass on scholarships because their English scores are still below standard. Mastery of English, especially the TOEFL score level, is important not only for getting scholarships or studying at certain universities but also sometimes a requirement for applying for jobs and getting promoted. The ability to speak English has significant meaning because it is very calculated in the world of education and the

world of work so that many improve their English language skills (Linda et al., 2021; Puspita & Hasyim, 2021; Sihite et al., 2021).

For this reason, it is important to increase the TOEFL score to equip people to face the global world with a very tight level of competition. Students are among the elements of society that must receive special attention in increasing mastery of foreign languages. Because students will face many possibilities after graduating from college, working, or continuing their studies to a higher level, which requires mastery of the English language, on this basis, it is urgent to train efforts to increase TOEFL scores for students.

Efforts made by tertiary institutions, especially Universitas Graha Nusantara, to improve English language skills, among others, by creating policies as a graduation requirement and implementing English language proficiency training programs, and carrying out English language tests in which the score of students' English skills is a reflection of their language skills. At Universitas Graha Nusantara, the standard for graduating students is English proficiency, with a minimum score of 400. The English language proficiency test at Universitas Graha Nusantara is conducted at the FKIP Language Laboratory; this language proficiency test is called the Institutional Toefl Based Test (ITBT).

Students are given intensive and consistent ITBT training for a week. Providing TOEFL materials or training for students will be very useful because it will support communication skills in English, especially in writing. The ITBT contains Listening, Structure, and Reading material. With Listening, students can have the ability to understand audio texts. Whereas in Structure material, students will be able to understand good and correct English, especially in writing. In Reading material, students will be able to understand a broader range of English because in Reading material, there are various discussions or themes. The implementation of training and tests conducted at the UGN FKIP Language Laboratory still uses paper-based tests because the TOEFL test requirements at several universities in Indonesia still use paper-based tests (Candria, n.d.; Sirajuddin & Yahrif, 2021).

A good training needs to be well-designed and evaluated for effectiveness to ensure that the training objectives are achieved (Rossi et al., 2018). Evaluation can mean providing information that can be presented as material for consideration in making a decision (Gusti & Ambiyar, 2021). By conducting an evaluation, weaknesses and problems that occur during the implementation of the training can be identified, and this information can be used as feedback in

planning future training. Various evaluation models have been developed and are very popular and widely used as strategies or work guidelines in implementing program evaluation; and each model has different characteristics (Ambiyar & Dewi, 2019), one of which is the Kirkpatrick Model (Kirkpatrick Four Levels Evaluation Model).

This study focuses on the Evaluation of the Kirkpatrick Model in terms of its goals in the field of education, some evaluations are macro, and some are micro (Purbadi, 2020; Sewang & Halik, 2019). A macro-targeted evaluation is an educational program, namely a program planned to improve the education sector. Micro-evaluation is often used at the class level, especially to determine student learning achievement. This learning achievement is not only cognitive but also includes all the potential in students. So, the target of micro-evaluation is the learning program in class, and the person in charge is the lecturer (Purwanta et al., 2020; Setyono, 2019).

The evaluation model developed by Kirkpatrick is known as the Kirkpatrick Four Levels Evaluation Model. According to Kirkpatrick, evaluation of the effectiveness of training programs (training) includes four levels of evaluation: level 1 reaction, level 2 learning, level 3 behavior, and level 4 result (Kirkpatrick & Kirkpatrick, 2006).

Evaluation of level 1 on the reactions of training/program participants means measuring participant satisfaction (customer satisfaction). The training program is effective if the training process is fun and satisfying for the trainees so they are interested and motivated to learn and practice. In other words, the training participants will be motivated if the training process runs satisfactorily for the participants, which will bring about pleasant reactions from the participants. Conversely, if the participants are not satisfied with the training process they are participating in, they will not be motivated to take part in further training.

At evaluation level 2, Kirkpatrick suggested three things an instructor could teach in a training program: knowledge, attitudes, and skills. Training participants are said to have learned if they have experienced a change in attitude, improved knowledge, or improved skills.

Evaluation at level 3 (behavioral evaluation) differs from evaluating attitudes at level 2. Attitude evaluation at level 2 focuses on changes in attitude that occur when training activities are carried out so that they are more internal. In contrast, behavior assessment focuses on behavior changes after participants return to the workplace. The success criterion for the 3rd level evaluation is that the change in attitude that has occurred after attending the training will be implemented

after the student attends the lecture so that the assessment of this behavior is more external. Evaluation of results at level 4 is focused on the final results that occur because participants have participated in a program. In learning activities, this evaluation model leads to the final results obtained by students.

Evaluation is carried out to obtain an overview or information about students' results and learning process, as well as to find out the difficulties that arise during the teaching and learning process. Evaluation is also closely related to the assessment process, which includes both the assessment of learning outcomes and the assessment of the process of learning. Evaluation is used to create a learning program that includes designs for teaching and learning. The evaluation also decides the position of a learning program based on certain criteria, so that a program can be trusted, trusted, and carried out continuously, or vice versa; the program must be fixed or improved (Glerum & Judge, 2021).

Kirkpatrick's model was chosen because it has been widely used to evaluate training programs around the world (Cutting et al., 2021, 2022; Siminto et al., 2021; Tsui, 2017). Therefore, this study uses the Kirkpatrick model in evaluating the Proficiency Test Preparation Course Program at the Language Laboratory of FKIP-UGN because it has advantages, which are comprehensive, simple, and applicable. Comprehensive means that this evaluation model is able to reach all sides of a training program. It is said to be simple because this model has a simple and easy-to-understand logic flow as well as clear and straightforward categorization. Meanwhile, in terms of usage, this model can be used to evaluate various types of training in various situations.

Currently, the Language Laboratory of FKIP UGN is more focused on designing, implementing training, and carrying out ITBT tests but still needs to conduct an adequate training evaluation process. This study aims to evaluate training organized by the UGN FKIP Language Laboratory based on the Kirkpatrick concept.

### *1.2. Research questions*

Based on the explanation from the background above, the research questions in this study are: How to evaluate the Proficiency Test Preparation Course Program at the Language Laboratory of FKIP UGN using the four-level Kirkpatrick model (Reaction, Learning, Behavior, and Result)?

### *1.3. Significance of the study*

This study aims to evaluate training organized by the Language Laboratory of FKIP UGN based on Kirkpatrick's concept so that it is known how effective the Proficiency Test Preparation

Course Program is in increasing students' language competency scores. Each aspect or level in the evaluation of the Kirkpatrick model is carried out with a different assessment or assessment model. This is done to get a general description of the abilities possessed by students in each competency on the ITBT test, as well as development and improvement for the implementation of the next program at the the Language Laboratory of FKIP UGN.

## **2. METHOD**

### *2.1. Research Design*

The approach used in this study is descriptive quantitative research with the type of research used evaluative research. Weiss in Sugiyono (2017) argues that evaluation research is applied research, a systematic way to determine the effectiveness of a program, action, policy, or other object being researched when compared to the objectives or standards applied. Evaluation research is carried out to improve the effectiveness of a program based on feedback from the people involved in implementing the program. Activities in evaluation research are comparing planned activities with activities implemented, comparing program objectives with the results achieved, and the positive or negative impacts of the program.

This research was conducted to evaluate the ITBT training management program, maintain the implemented program, and improve program quality. The evaluation model used is the Kirkpartric model with 4 levels: Reaction, Learning, Behavior, and Result.

### *2.2. Samples/Participants*

Retrieval of research subjects or data sources to be used as sources of information in this research must be adjusted to the goals and information needs of researchers. Therefore, in this study, the selection of research subjects used a "purposive sampling" technique, namely the collection of data sources/subjects based on research choices about what aspects and whom to focus on during certain situations and at this time continuously throughout the research. Purposive sampling depends on the purpose of the focus at the time of research (Nasution, 2003). Based on the explanation above, the subjects of this study were 30 students who attended language training from various study programs at Universitas Graha Nusantara.

### 2.3. Instruments

Data collection in this study was carried out for each aspect of the Kirkpatrick Model evaluation. The following are the instruments used in this study at each evaluation level of the Kirkpatrick Model.

**Table 1. Instruments each level of the Kirkpatrick Model evaluation**

<b>Level</b>	<b>Measurement</b>	<b>Instruments</b>
<b>Reaction</b>	Measuring the level of satisfaction of the training participants with the training programs they participated in	Questionnaire and the Interval Measurement Scale (scale Likert)
<b>Learning</b>	Measure the level of learning followed	formal test participants at each meeting
<b>Behavior</b>	Measuring changes in student behavior towards the benefits of language skills training provided	Questionnaire and the Interval Measurement Scale (scale Likert)
<b>Results</b>	Measuring final student results	Formal test: pre-test and post test

### 2.4. Data analysis

Analysis of research data was carried out using SPSS and Microsoft Excel software. Data obtained from various data collection techniques were then analyzed. Analysis of the data obtained was carried out using the steps proposed (Nugrahani, 2014) including (1) preparation, (2) classification, (3) processing/processing, and (4) interpretation and conclusion.

To determine the level of success of the research conducted, it is necessary to identify the scores obtained. The formula used to find the percentage (Sudijono, 2016). After completing the data collection, the researcher tried to compile and classify the data and select the data in this study. This serves as an answer to the formulation of the problem that has been determined. According to (Arikunto, 2006), the data obtained in this study were

**Table 2. Success Criteria**

<b>Score Criteria</b>	<b>Score</b>
Very Satisfied	76-100%
Satisfied	56-75%
Partly Satisfied	41-55%
Not at all Satisfied	0-40%

### **3.FINDINGS AND DISCUSSION**

#### *3.1. Findings*

The evaluation implementation of the Kirkpatrick model has been carried out in the Proficiency Test Preparation Course Program in the Language Laboratory. This training is given to students who wish to take the ITBT test as a prerequisite for the final undergraduate exam, with a minimum score of 400 for non-English Program students and a score of 450 for students of English study programs. The training time is 4 days; before the training, students are given a pre-test, and after the training, they are given a post-test. The results of data analysis on the implementation of the Proficiency Test Preparation Course Program used the evaluation of the Kirkpatrick Model, which was explained in four evaluation aspects. The data obtained at each evaluation level uses a different assessment.

#### **Evaluating Reaction**

A learning program is considered successful if students are satisfied with the strategy for delivering the material, learning media, and assessments used. Harackiewicz et al. (2016) suggests the success of a training program depends on how interested, focused, and motivated the participants are. People learn better when they react positively to the learning environment. It was concluded that the success of the process of learning activities is inseparable from the interest, attention, and motivation of students in following the course of lectures. Students learn better when they react positively to the learning environment. Student satisfaction with the learning process carried out is examined from several aspects, namely Training Materials, Training Methods, Training Instructors, and Training facilities. On the training material aspect, there are 4 questions, on the training methods aspect, there are 3 questions; on the training instructor aspect there are 7 questions, and on the training facilities aspect there are 6 questions.

Before the questionnaire was distributed to the subjects, validity was previously carried out to ascertain how well an instrument was used to measure the concept that should be measured by Sugiyono (2017). Each instrument item from the reaction evaluation obtained  $r_{count} > r_{table}$  (0.514), then the instrument from the evaluation reaction was declared valid. The questionnaire is reliable if the reliability coefficient is positive and greater than 0.60. As for the results of the reliability test on the reaction evaluation instrument, it is known that each variable has a Cronbach Alpha of 0.874 more than 0.60 ( $\alpha > 60$ ), so it can be concluded that the reaction evaluation instrument is reliable. Next, the distribution of questionnaires was carried out on the subject, namely the training participants.



**Table 3. Evaluating Student Reactions to The Implementation of Training**

<b>Aspect</b>	<b>Percentage</b>	<b>Description</b>
Training Materials	88.50%	Very Satisfied
Training Methods	88.00%	Very Satisfied
Training Instructor	86.19%	Very Satisfied
Training facilities	85.56%	Very Satisfied
<b>Mean</b>	<b>87.06%</b>	<b>Very Satisfied</b>

Table 3 shows that overall, the average percentage of student responses to the evaluation reactions of the Proficiency Test Preparation Course Program is 87.06%, included in the Very Satisfied category. It means that the training process is felt satisfying, fun, and beneficial for the participants.

Perceptions of students regarding **training materials** have a percentage of 88.50% and are included in the **Very Satisfied** category. Students feel that the training material provided can be valuable and exciting and is discussed in depth by the instructor. Perceptions of training participants regarding **training methods** have a percentage value of 88.00%. They are included in the **Very Satisfied** category, meaning that the Proficiency Test Preparation Course Program uses training methods that follow the topics discussed. The training is carried out according to the schedule determined by the training committee and on time. The training committee is also alert in providing assistance to students. The perception of the training participants regarding the **training instructor** has a percentage value of 86.19% and is included in the **Very Satisfied** category. It means that the instructor can convey the training material clearly and well so that the participants can understand the training material provided.

Student perceptions of **training facilities** have a percentage value of 85.56% and are included in the **Very Satisfied** category. Students are satisfied with the facilities provided in training; the participants feel comfortable with clean and good facilities. The facilities used during the training are complete with supporting the implementation of the training, such as LCD projectors, office stationery, notes, and training modules provided to the training participants to help them learn and understand the training material.

From the results obtained in table 3, it can be said that students who take part in the Proficiency Test Preparation Course Program have a positive reaction to the learning process being carried out. This evaluation is a reference for making improvements or improving the quality of the teaching and learning process by lecturers.

## Evaluation Learning

Measuring the effectiveness of learning programs is necessary to carry out learning evaluations (Evaluating Learning). Evaluation of learning involves three aspects: changing attitudes, increasing knowledge, and improving skills in students. Learning will only be considered a success if these three aspects are carried out or if there is change. Evaluating learning is more directed at evaluating learning results (output). So that in the assessment of learning outcomes (learning measurement) performance assessment is used to measure the knowledge that has been learned, changes in attitudes, and skills that have been developed or improved.

The participating students were 30 students, consisting of 6 students in the State Administration study program, 4 students in the Agrotechnology study program, 1 student in Development Economics study program, 7 students in Government Science study program, 7 students in Management study program, Animal Husbandry with 4 students, and Civil Engineering study program with 1 student.

The time and duration of the program is 4 days for one week, and applied in 1 hour 30 minutes per day. The learning location is in the language laboratory of FKIP UGN. The Proficiency Test Preparation Course Program is quite good and smooth. Students go through all the learning processes by referring to the teaching materials according to the modules provided by the trainer.

The material is divided into three groups: general ITBT information, discussion of ITBT questions, and simulations. In part 1, which is listening, there are 50 questions of three different types: short conversations, long conversations, and long stories (Long Talk). In the second part, there are two kinds of questions: those that look like structures (Sentence Completion) and those that look like written expressions (Error Identification). There are 40 questions in each type. Students have to choose the right answer to multiple-choice questions about structure in the structure section. In the English structure of the completion section, students are told to look at tenses, prepositions, and spellings in terms of grammar. And the third is Reading Comprehension, which has up to 50 questions about vocabulary and reading passages. A post-test is done after the pre-test and the core activities (training). Here is the average score on the pre-test for each of the 30 students who took it.

**Table 4. Pre-test**

Aspect	Mean
LC	36.9
S&WT	36.7
RC	37

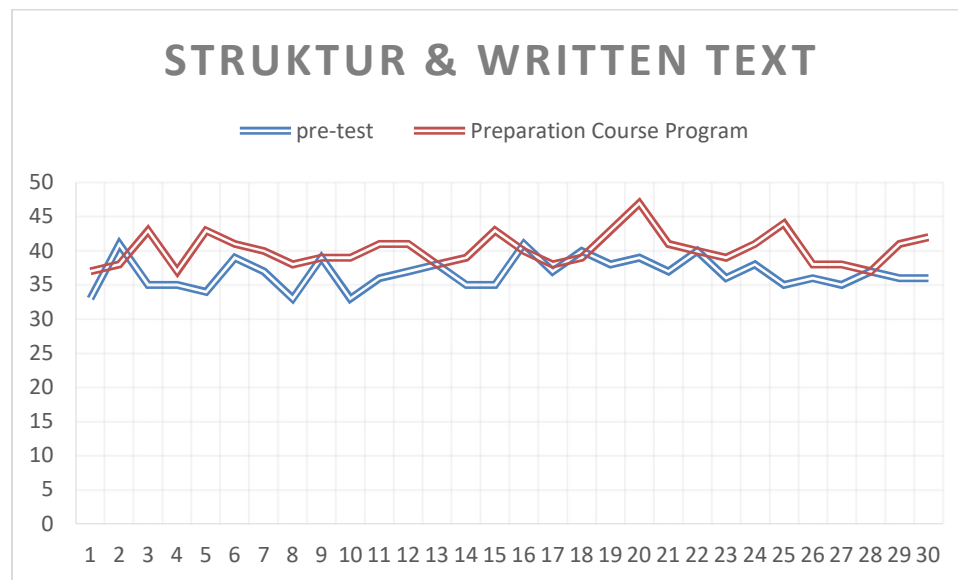
Table 4 shows that the average student has the lowest score in Structure & Written Expression with an average score of 36.7, then in Listening comprehension with an average score of 36.9 and reading comprehension with an average score of 37.

According to the training plan, this training will focus on part 2, which has the lowest score (Structure & Written text). The low score comes from not being able to write correctly and properly. According to Amilia (2018), writing skills are skills that require complex language skills and a high level of competence. At the very least, you need to be able to read and listen in order to write well. Like speaking, writing is a skill that can help you do things. Writing, on the other hand, is thought to be harder than speaking. This is because writing activities have a lot of rules to follow.

The following is the increase in the results of Structure & Written Text during the training:

**Table 5. The Results of Structure & Written Text during the training**

	Pre-test	Preparation Course Program	Improvement
Mean	36.70	40.23	3.47
Percentage	54%	59%	5%



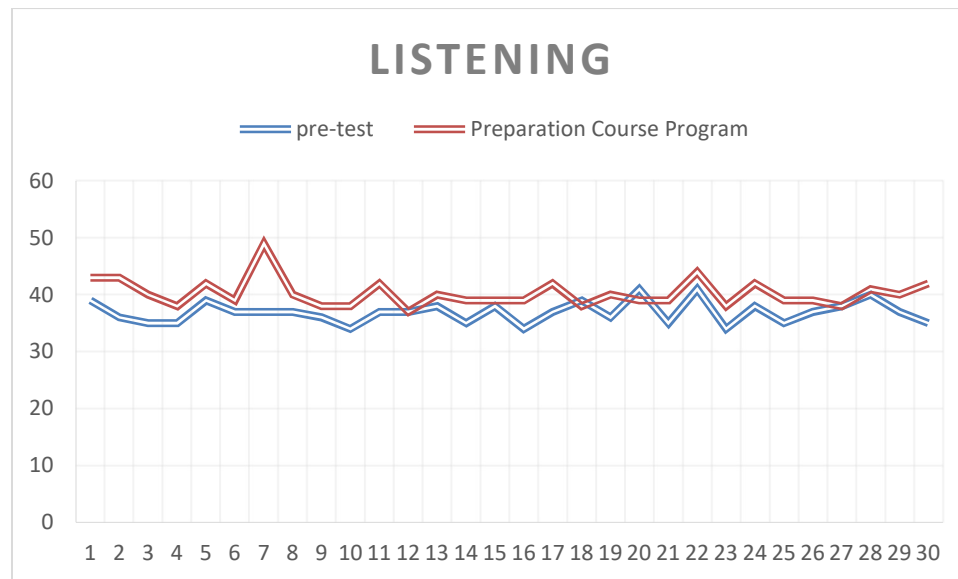
**Figure 1. The Result of Structure & Written Text during the training**

Table 5 above shows that there was an increase in student scores on the Structure & Written Text aspect, which was higher than the pre-test. The average value of Structure & Written text increased from 36.70 in the pre-test to 40.23 in training. There was an increase of 5 percent compared to the pre-test. Tutors said they still emphasize Grammar and Writing because of the difficulties faced by students, and the achievements obtained are lower than in Listening Comprehension and Reading Comprehension. The training test results showed an increase in students' Structure & Written Text scores.

The following is the increase in the results of Listening Comprehension during the training:

**Table 6. The Result of Listening Comprehension during the training**

	Pre-test	Preparation Course Program	Improvement
Mean	36.90	40.20	3.33
Percentage	54%	59%	5%



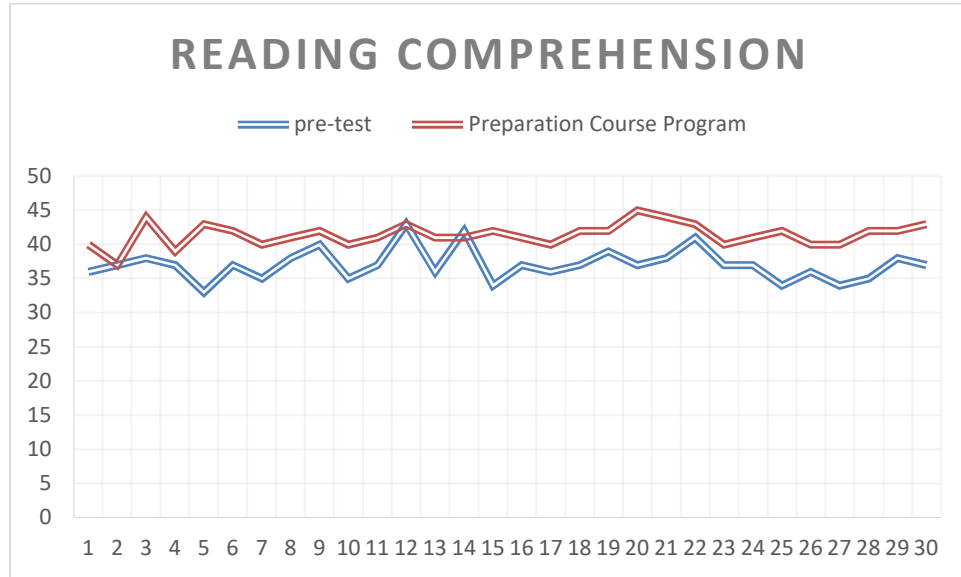
**Figure 2. The Result of Listening Comprehension during the training**

Table 6 above shows an increase in students' scores on the Listening Comprehension aspect, which is higher than the pre-test. The average value of Listening Comprehension increased from 36.9 in the pre-test to 40.20 in training. There was an increase of 5 percent compared to the pre-test. The training test results showed an increase in students' Listening Comprehension scores.

The following is the increase in the results of Reading Comprehension during the training:

**Table 7. The Result of Reading Comprehension during the training**

	Pre-test	Preparation Course Program	Improvement
Mean	37.03	41.43	3.33
Percentage	55%	62%	7%



**Figure 3. The Result of Reading Comprehension during the training**

Table 7 above shows an increase in students' scores on the Reading Comprehension aspect, which was higher than the pre-test. The average Reading Comprehension score increased from 37.03 in the pre-test to 41.43 in training. There was an increase of 7 percent compared to the pre-test. The training test results showed an increase in students' reading scores.

### Evaluation Behavior

Evaluation of behavior related to the Proficiency Test Preparation Course Program process emphasises changing student behaviour towards the benefits of the training provided. This evaluation is likely successful if students can implement the Proficiency Test based on the phenomena of everyday life.

At this behaviour stage, participants were asked to describe the form of behavior change through answers to open interviews regarding: Actively asking questions to develop knowledge,

the effect of the Proficiency Test Preparation Course Program on performance achievement, and the ability to help colleagues. Each of them contains 2 question items.

**Table 8. Evaluation of Student Behavior on the Implementation of Training**

<b>Aspect</b>	<b>Total</b>	<b>Percentage</b>	<b>Description</b>
Actively ask questions to develop knowledge	255	85.00%	excellent
influence of the Proficiency Test Preparation Course Program on performance achievement	247	82.33%	excellent
ability to help colleagues	257	85.67%	excellent
<b>Mean</b>		<b>84.33%</b>	<b>excellent</b>

Based on table 8, it can be seen that overall the average percentage of students' responses to the Proficiency Test Preparation Course Program, which includes the overall Evaluation Behavior is 84.33%, included in the excellent category, meaning that students are active during training because they realize the need for language skills to the workplace and willing to help other students in understanding the material.

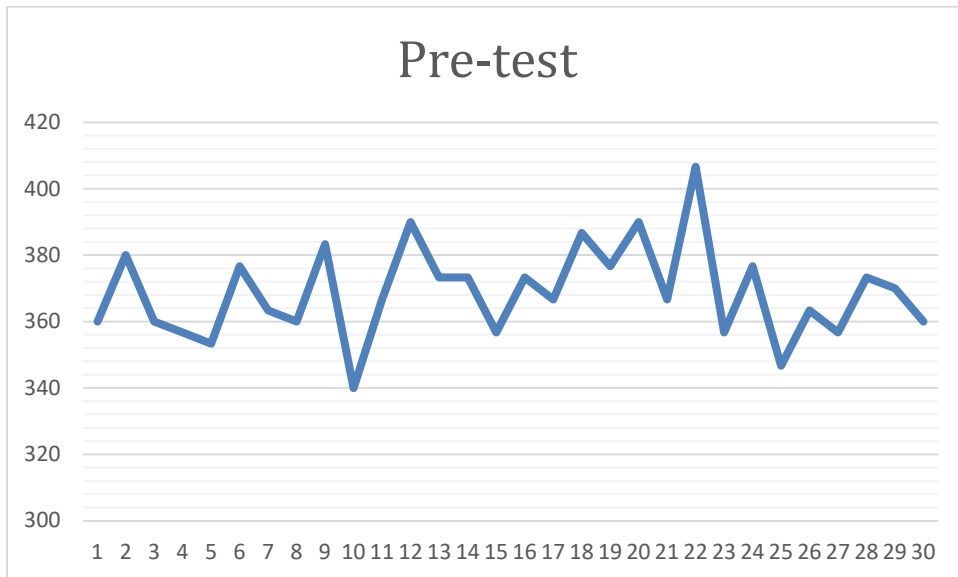
### **Evaluation Result**

Evaluation of results is related to the final results of students. This evaluation is likely successful if students are able to achieve good academic scores.

This activity is one of the ITBT training series that determines the success or failure of this training. The success of the ITBT training was marked by an increase in the participant's ITBT score from the previous test, namely the pre-test. If there is an increase, the post-test results are greater than the pre-test, and this training is considered successful. Conversely, if the average participant has decreased, the training is unsuccessful. Below is the ITBT training post-test result

**Table 9. Empirical Data Categorization of Pre-Test Scores**

<b>Category</b>	<b>Criteria</b>	<b>Frequency</b>	<b>Percentage</b>
Low	< 355	3	10.00
Medium	355 - 383	22	73.33
High	383 <	5	16.67
<b>Total</b>		<b>30</b>	<b>100.00</b>

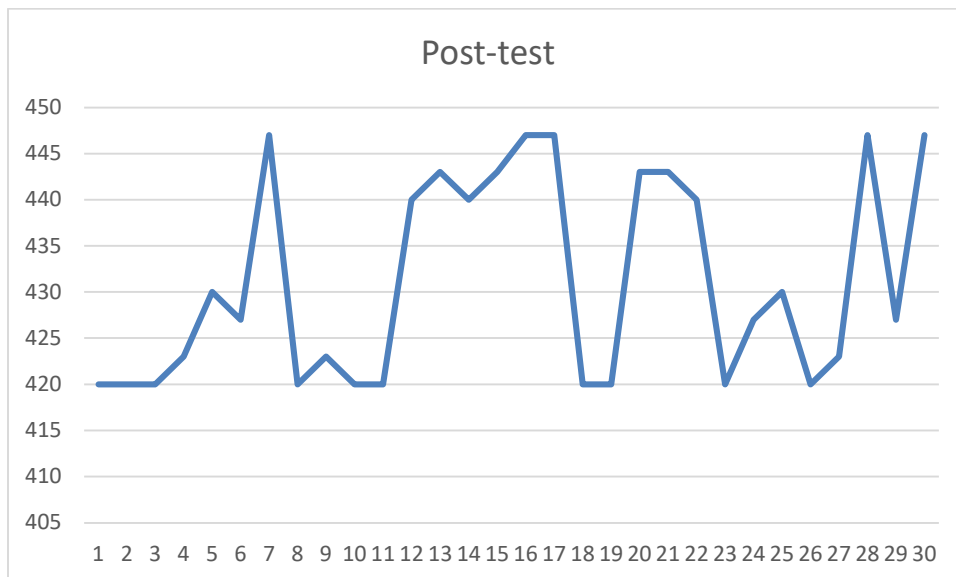


**Figure 4. The Result of Pre-test**

Table 9 above shows that the average score of students is still in the medium category of 22 students and in the low category of 3 students. This indicates that the ITBT score is still far from the standard, which is 400. There are only 5 participating students who have scores above 400 with a score range of more than 383.

**Table 10. Empirical Data Categorization of Post Test Scores**

Category	Criteria	Frequency	Percentage
Low	< 420	0	0.00
Medium	420 - 442	21	70.00
High	442 <	9	30.00
<b>Total</b>		<b>30</b>	<b>100.00</b>



**Figure 5. The Result of Post-test**

Based on table 10 of the empirical data categorization above, it can be seen that the average score of students in the medium category with criteria 420-442 is 21 students, and in the high category there are 9 students with criteria more than 442. This indicates that the students' ITBT scores in the post-test have met the above standard of 400.

### 3.2. Discussion

At level 1 of the reaction evaluation, it shows that the overall average percentage of student responses to the Proficiency Test Preparation Course Program is 87.06%, included in the Very Satisfied category, meaning that the training process is felt satisfying, fun and useful for the trainees.

The high satisfaction of the training participants will affect the increase in their knowledge of the training participants. Training participants who are highly satisfied with the training will learn the material presented during the training well. In this study, student perceptions stated that they were very satisfied with a percentage of 88.50% in the training material delivered by the ITBT Trainer in the Language Laboratory FKIP UGN. There is a relationship between training participant satisfaction and the learning level of trainees in accordance with the results of research conducted Nuraini (2018) that the more competent the trainer, the higher the quality of service will make students feel satisfied and improve learning outcomes. Feigenbaum (1997) stated that quality is full of customer satisfaction.



According to research conducted by Solikha & Rasyida (2020), the success of a training program in increasing the knowledge, skills and attitudes of students depends on the content and methods of training. According to research conducted by (Ooi, 2007), training methods also contribute to achieving training effectiveness. This is in line with the results of this study; at the reaction evaluation level, 88.00% of students stated that they were very satisfied with the Training Method delivered by the ITBT Trainer at the Language Laboratory FKIP UGN.

In terms of training instructors received Very Satisfied responses from students with a percentage of 86.19%. Training methods and materials that attract the attention and interest of the training participants are supported by training trainers who understand the training material and answer participants' questions well, which also supports the achievement of training effectiveness. The research conducted Haslinda & Mahyuddin (2009) states that a significant factor contributing to training effectiveness is the trainer's (instructor) competence.

The level of student satisfaction in the comfort of the classroom used for training with very satisfying criteria with a percentage value of 85.56%. Noe & Kodwani (2018) explained that the room for training must be 1) comfortable and easily accessible, 2) quiet, private, and free from distractions, 3) wide for participants to move around. Sari & Mundzir (2017) in their research found that a hot and uncomfortable room was why the participants were passively involved during the training.

At level 2 of evaluation learning, there is an increase in students' abilities in listening competence, structure and written tests, and reading comprehension. This is due to the instructor's ability to convey material clearly at the active listening level. Based on research conducted by Noviyenty (2018), there was a significant increase in students' ITBT scores after they attended the ITBT Training Program at the Language Center. Scores of each aspect of ITBT also increased significantly.

Learning can be defined as a change in attitude, improvement, knowledge and or skill increase of participants after completing a training program. Training participants are said to have learned if they have experienced a change in attitude, improved knowledge, or increased skills (Nurhayati, 2018).

In this study level 3 behavior evaluation, overall, the average percentage of students' responses to the Proficiency Test Preparation Course Program was 84.33%, included in the

excellent category, meaning that students were active during the training because they realized the need for language skills for the world of work and were willing to help other students in understanding the material. The behaviorist theory introduced by Ivan Pavlov and developed by Thorndike and Skinner (in Fathurrohman, 2017) argues that learning is related to changes in behavior. Teaching and learning will affect all actions or behavior of a person so that thus a person's behavior can be noticed and predicted. Based on Kirkpatrick's theory, training participants are said to learn if they have experienced increased knowledge (Kirkpatrick & Kirkpatrick, 2006). Increased knowledge that occurs in training participants in understanding the training material.

Evaluation of results in the 4th level is focused on the final results that occur because students have attended the Proficiency Test Preparation Course Program. The final outcome category of a learning program includes increasing learning outcomes, knowledge, and skills (Cross et al., 2010). This becomes a reference for the application of evaluation results by looking at the final results with the expected achievement categories. The final success rate of the evaluation results is based on the empirical data categorization table average student score in the medium category with criteria 420-442 is 21 students, and in the high category there are 9 students with criteria more than 442. This indicates that the students' ITBT scores in the post-test have met the above standard of 400.

The evaluation of the Kirkpatrick model in this study is to measure how far the students' foreign language skills affect the achievement of the Proficiency Test Preparation Course Program targets, which affect students' achievement of their language competence. It is understood that the Proficiency Test Preparation Course Program emphasizes knowledge and the attitudes and skills of the participants, in general, to increase their competence, specifically in facing the challenges of globalization. Based on the evaluation findings that have been put forward, it can be seen that the impact of the Proficiency Test Preparation Course Program at the Language Laboratory of FKIP UGNP is clear that there is an increase in the performance of the alum participants in better job systematics.

This research was conducted by following the stages or levels in evaluating the Kirkpatrick model. Still, there were some research limitations, including the number of samples or research classes consisting of one class. The observation process is carried out in every rubric aspect, and the number of sampled students is one of the limiting factors in this study. It is hoped that further research can be used in a wider scope in terms of population and sample size.

## 4. CONCLUSIONS

### 4.1. Conclusion

This study aims to evaluate the implementation of the Proficiency Test Preparation Course Program at the Language Laboratory of FKIP UGN using the Kirkpatrick evaluation model because this model has never been evaluated so far. This study shows that the four variables in Kirkpatrick's model can be applied well to evaluate the implementation of the Proficiency Test Preparation Course Program. Differences in educational background and length of program participants did not significantly affect the evaluation results. The overall evaluation results show that the Proficiency Test Preparation Course Program implemented at the Language Laboratory of FKIP UGN is good.

Based on the evaluation results, the Proficiency Test Preparation Course Program at the Language Laboratory of FKIP UGN, can be continued because it helps non-English Department students develop foreign language competence. The results of this evaluation can be used as material for consideration in determining steps that can be taken to improve the implementation of the Proficiency Test Preparation Course Program so that the results of this program can be achieved optimally.

### 4.2. Suggestions

In order to improve the quality of organizing the Proficiency Test Preparation Course Program at the Language Laboratory of FKIP UGN, it is necessary to monitor evaluation in post-training evaluation for program development and program sustainability. The evaluation method used in this study is the Kirkpatrick method. For future researchers interested in conducting similar research, it is recommended to use other evaluation methods that follow the objectives of implementing the program.

## REFERENCES

- Ambiyar, A., & Dewi, M. (2019). *Metodologi Penelitian Evaluasi Program*. Alfabeta.
- Amilia, F. (2018). Pemahaman dan habituasi untuk membangun kompetensi menulis praktis dan ilmiah. *Lingua Franca: Jurnal Bahasa, Sastra, Dan Pengajarannya*, 2(1). <https://doi.org/http://dx.doi.org/10.30651/lf.v2i1.1401>
- Arikunto, S. (2006). *Prosedur Penelitian Tindakan Kelas*. Bumi Aksara.
- Candria, M. (n.d.). Pelatihan Paper-Based TOEFL Bagi Mahasiswa Fakultas Syariah Dan Hukum, Universitas Islam Negeri (UIN) Walisongo, Semarang. *Harmoni: Jurnal Pengabdian Kepada Masyarakat*, 6(1), 93–97. <https://doi.org/10.14710/hm.6.1.93-97>
- Cutting, M., Goda, Y., & Suzuki, K. (2021). Analysis of Behaviors Documented in Reflective Journals in Study Abroad ePortfolios to Improve Pre-Departure Training. *International*

- Journal*, 15(1), 117–128. <https://ijemt.org/index.php/journal/article/view/23>
- Cutting, M., Goda, Y., & Suzuki, K. (2022). Evaluation of Pre-Departure Training for Study Abroad by Adapting Kirkpatrick's Model. *International Journal for Educational Media and Technology*, 16(1), 27–37. <https://ijemt.org/index.php/journal/article/view/272>
- Fathurrohman, M. (2017). *Belajar dan pembelajaran modern: konsep dasar, inovasi dan teori pembelajaran*. Garudhawaca.
- Feigenbaum, A. V. (1997). Changing concepts and management of quality worldwide. *Quality Progress*, 30(12), 45.
- Glerum, D. R., & Judge, T. A. (2021). Advancing employability: applying training evaluation to employability development programs. *Career Development International*, 6(3), 363–390. <https://doi.org/10.1108/CDI-09-2020-0248>
- Gusti, S. M., & Ambiyar, A. (2021). Evaluasi Layanan Bimbingan Teknologi Informasi dan Komunikasi (TIK). *Jurnal Edutech Undiksha*, 9(2), 174–180. <https://doi.org/10.23887/jeu.v9i2.38949>
- Harackiewicz, J. M., Smith, J. L., & Priniski, S. J. (2016). Interest Matters: The Importance of Promoting Interest in Education. *Policy Insights from the Behavioral and Brain Sciences*, 3(2), 220–227. <https://doi.org/10.1177/2372732216655542>
- Haslinda, A., & Mahyuddin, M. Y. (2009). The effectiveness of training in the public service. *American Journal of Scientific Research*, 6(1), 39–51.
- Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating Training Programs: The Four Levels*. Berrett-Koehler Publishers, Inc.
- Linda, D., Hasra, H., & Anindhita, W. (2021). Penguatan Keterampilan Berbahasa Prajurit TNI AL di KRI Sutanto-377 Melalui Modul Pembelajaran Bahasa Inggris untuk Mewujudkan TNI AL Berkelas Dunia (World Class Navy). *Prosiding Seminar Nasional Bahasa, Sastra, Dan Seni*, 1, 168–171.
- Nasution, S. (2003). *Metode Research (penelitian ilmiah)*. Bumi Aksara.
- Noe, R. A., & Kodwani, A. D. (2018). *Employee training and development, 7e*. McGraw-Hill Education.
- Noviyenty, L. (2018). An evaluation of TOEFL matriculation program for STAIN students. *English and Literature Journal*, 5(1), 55–68. <https://doi.org/10.24252/elite.v5i1a6>
- Nugrahani, F. (2014). *Metode Penelitian Kualitatif dalam Bidang Pendidikan Bahasa*. Cakra Books.
- Nuraini, N. (2018). Tingkat Kepuasan Peserta Terhadap Kualitas Pelayanan Widyaiswara Pada Diklat Teknis Substantif Publikasi Ilmiah bagi Guru Mata Pelajaran Agama MTs Angkatan III. *Andragogi: Jurnal Diklat Teknis Pendidikan Dan Keagamaan*, 6(1), 168–186. <https://doi.org/10.36052/andragogi.v6i1.53>
- Nurhayati, Y. (2018). Penerapan Model Kirkpatrick untuk Evaluasi Program Diklat Teknis Substantif Materi Perencanaan Pembelajaran Di Wilayah Kerja Provinsi Kepulauan Riau. *Andragogi: Jurnal Diklat Teknis Pendidikan Dan Keagamaan*, 6(2), 170–187. <https://doi.org/10.36052/andragogi.v6i2.63>
- Ooi, A.-Y. (2007). The determinants of training effectiveness in Malaysian organizations. *International Journal of Business Research*, 12(3), 321–333.
- Purbadi, P. (2020). Manajemen Pembelajaran Berbasis Sekolah di SMK Muhammadiyah 1 Playen Gunungkidul. *Media Manajemen Pendidikan*, 3(2), 232–241. <https://doi.org/10.30738/mmp.v3i2.4663>
- Purwanta, W., Syahrial, Z., Wibawa, B., & Siang, J. L. (2020). Training Evaluation Program for the Development and Empowerment Center of Educators and Language Education Personnel in 2019 (PPP in the Ministry of Education and Culture). *International Journal of Advanced Science and Technology*, 29(4s), 1694–1702.

- <https://confbrite.org/2020/smic/kfz/abstract/Ck9Pj6art>
- Puspita, N., & Hasyim, U. A. (2021). Pendampingan Kegiatan English Club Mahasantri Pondok Pesantren Di Kota Metro Melalui Program Peningkatan Kemampuan English Proficiency (Toefl). *Wisanggeni: Jurnal Pengabdian Masyarakat*, 23–31. <https://doi.org/10.25217/wisanggeni.v1i2.1972>
- Rofaida, R., & Gautama, B. P. (2019). Strategi Peningkatan Kompetensi Lulusan Perguruan Tinggi melalui Studi Pelacakan Alumni (Tracer Study). *Image: Jurnal Riset Manajemen*, 8(1), 1–8. <https://doi.org/10.17509/image.v7i1.23171>
- Rossi, P. H., Lipsey, M. W., & Henry, G. T. (2018). *Evaluation: A systematic approach*. Sage publications.
- Sari, R. W., & Mundzir, S. (2017). Keterlibatan Peserta Pada Proses Pelaksanaan Pelatihan Subkejuruan Mobil Diesel di Unit Pelaksana Teknis Pelatihan Kerja (UPT-PK) Singosari Malang. *Ilmu Pendidikan: Jurnal Kajian Teori Dan Praktik Kependidikan*, 2(1), 76–84. <https://doi.org/10.17977/um027v2i12017p076>
- Setyono, B. (2019). Micro-evaluation of Writing Materials Designed by Student Teachers of Language Education: The Perspective of Genre Pedagogy. *The 63rd Teflin International Conference 2016 Proceeding: Creativity And Innovation In Language Materials Development And Language Teaching Methodology In Asia And Beyond*, 266–273.
- Sewang, A., & Halik, A. (2019). Model Manajemen Pembelajaran Pendidikan Islam Berbasis Masalah: Studi Kasus pada Jurusan Tarbiyah dan Adab IAIN Parepare. *JPPI (Jurnal Pendidikan Islam Pendekatan Interdisipliner)*, 3(1), 1–15.
- Sihite, H. Y., Setyanto, H. Y., & Gervinko, G. (2021). Analisis dan Strategi Pengembangan Kemampuan Bahasa Inggris Pariwisata di SMAN 1 Teluk Sebong Kab. Bintan. *Jurnal Konsepsi*, 10(3), 269–276. <https://www.p3i.my.id/index.php/konsepsi/article/view/120>
- Siminto, S., Soesanto, S., Sutarto, J., & Handoyo, E. (2021). Foreign Language Strengthening Programs Characteristics For Students In Non-Formal Learning At Several Universities In Indonesia. *NVEO-Natural Volatiles & Essential Oils Journal*, 8(6), 1172–1185.
- Sirajuddin, S., & Yahrif, M. (2021). Pelatihan Listening Untuk TOEFL Test Bagi Mahasiswa Prodi Sarjana Pendidikan Bahasa Inggris. *MEGA PENA: Jurnal Pengabdian Kepada Masyarakat*, 1(1), 67–75. <https://doi.org/10.37289/mp.v1i1.24>
- Solikha, N., & Rasyida, I. (2020). Schoology Terhadap Peningkatan Keaktifan Dan Hasil Belajar Siswa X Ips Man Kota Pasuruan. *Jurnal Ilmiah Edukasi & Sosial*, 11(1), 31–42. <https://jiesjournal.com/index.php/jies/article/view/221>
- Sudijono, A. (2016). *Pengantar evaluasi pendidikan, Cetakan ke-15*. Rajagrafindo Persada (Rajawali Pers).
- Sugiyono. (2017). *metode penelitian kuantitatif, kualitatif, R & D*. CV Alfabeta.
- Thariq, P. A., Husna, A., Aulia, E., Djusfi, A. R., Lestari, R., Fahrimal, Y., & Jhoanda, R. (2021). Sosialisasi pentingnya menguasai bahasa Inggris bagi mahasiswa. *Jurnal Pengabdian Masyarakat: Darma Bakti Teuku Umar*, 2(2), 316–325. <https://doi.org/10.35308/baktiku.v2i2.2835>
- Tsui, C. (2017). EMI teacher development programs in Taiwan. In *English as a medium of instruction in higher education* (pp. 165–182). Springer.
- Ucu, K. R. (2021). Literasi Indonesia Ketinggalan Kereta. <https://www.Republika.Co.Id/>. <https://www.republika.co.id/berita/r0jsu5282/literasi-indonesia-ketinggalan-kereta>