

## Evaluation of TOEFL preparation course program to improve students' test score

Mega Selvi Maharani\*; Nur Hidayanto Pancoro Setyo Putro

Universitas Negeri Yogyakarta

Jl. Colombo No. 1, Karangmalang, Depok, Sleman, Yogyakarta 55281, Indonesia.

\*Corresponding Author. E-mail: megaselvi6@gmail.com

### ARTICLE INFO

#### Article History

**Submitted:**

12 March 2021

**Revised:**

10 May 2021

**Accepted:**

31 May 2021

#### Keywords

TOEFL preparation;  
course program;  
improving students' test  
score

#### Scan Me:



#### How to cite:

Maharani, M., & Putro, N. (2021). Evaluation of TOEFL preparation course program to improve students' test score. *Jurnal Penelitian dan Evaluasi Pendidikan*, 25(1), 63-76. doi:<https://doi.org/10.21831/pep.v25i1.39375>

### ABSTRACT

This research was conducted to evaluate the management program in English village, maintain the implemented program, and increase program quality. The method used in this study was descriptive qualitative and quantitative research. The model evaluation used was CIRO (Context, Input, Reaction, and Outcome) by War, Bird, and Rackman. The subject of this research were 30 students, four tutors, and a program director. Interview, survey, and observation were used for collecting the data. The analyzed quantitative data used descriptive percentages, and qualitative data used condensation, data exposure, conclusion drawing, and verification. The results of this study show that (1) context in this study considered to participant needed which synchronized to the standard implementation of the TOEFL real test and national standard. (2) input has been provided and is well prepared, including program plan, tutor qualification, for admission of course participants, and facilities. (3) Reaction of the participants' satisfaction was in the satisfactory category, and the participant's assessment reaction of the TOEFL program implementation process was in the very good category. (4) Outcome obtained by course participants in the TOEFL scoring has not increased significantly; it has even decreased during the last test.

This is an open access article under the **CC-BY-SA** license.



## INTRODUCTION

The development of education today has always been a top trending in every discussion; each sector generally contributes to educational progress. Efforts to improve education quality are made by many people using various strategies (Kurniawati, 2017). Educational quality in Indonesia is one of the objectives and integral pieces of developing human resources exhaustively (Mulyasa, 2005). Seeing the importance of improving the quality of education, it takes effort and cooperation that is actualized continuously by foundations or institutions of formal, informal, and non-formal education.

Non-formal education is an effort to improve the education provided by the community and given to the societies according to existing needs (Sudjana, 2004). Most of the non-formal education established in Indonesia is approximately 16,935, including courses and training. Most of them are English courses due to so many requests from the community (Direktorat Pembinaan Kursus dan Pelatihan, 2018). With many courses provided in Indonesia, the quality of English graduates should be better. On the contrary, Indonesia is the lowest rank of mastery of the English language in Asia. Dahuri (2019) states that Indonesia is one of the countries that are still included in the low English Proficiency category. The low level of English proficiency in Indonesia is caused by the poor environment of English (Maruf et al., 2020). It can be seen on the data of EF English Proficiency Index that Indonesian is included on low

rank which has 15th on Asian number and 453 test score (English First, 2020). One of the lousy environmental impacts is students' low motivation in learning English. Thus, in this case, the English village comes to build an environment that supports learning English by developing the communities and raising many institutions to improve students' English skills.

In the English village, many language communities have the same vision to learn English (Nurhayati et al., 2013). English village course institutions provide language services such as basic classes, intermediate classes, and advanced classes. The final program in the English class is the TOEFL preparation program. TOEFL preparation program is a learning program to improve student reading, structure, and listening skills in the academic English language (Sakurai, 2020). TOEFL is one of the world's English proficiency tests (Ismail & Othman, 2020; Syamsuddin & Min, 2014). Furthermore, according to Ali (2012), TOEFL is an accumulation of student learning outcomes and achievements in English. Thus, this clearly shows that the TOEFL test is a measure of students' successfully mastering English, so the TOEFL test preparation program is needed to improve the scores of participants in the test. According to Ma and Cheng (2016), taking a test preparation course is the most time-efficient method of preparing for the TOEFL test. TOEFL preparation has been held out in the Pare English village, but many problems are found in implementing TOEFL preparation courses.

Based on the pre-research in December 2019, the researchers found several problems in some courses in the Pare English village, one of which is the TOEFL program. The problem is that students' TOEFL scores had not reach the national standard yet. Students' TOEFL score in the final test was less than 450 score. However, the average Indonesian minimum standard score is 450 as the requirements of the college. In order to improve students' scores, the course plan must be well designed to help increase the scores of course participants (Pre-research data, 2019). Further, another study found that the problem that is often mistaken at holding in TOEFL courses is that the tutor spends much time providing strategies for answering questions to increase student scores, but they are ignoring participant comfort in the class so that participants feel that the TOEFL class is boring (Wang, 2019). The organizer must handle these problems so it can give an attractive impression to the students in the learning process without neglecting the actual course objectives. To overcome this case, it is necessary to have good management in a course institution. As such, for controlling the management program, an evaluation is a need when the program has been running. An evaluation is needed to see the effectiveness of a program since it is an activity to improve quality, performance, and productivity in implementing a program (Mardapi, 2017). Problems in a program can be overcome with an evaluation which is the basis for decision making (Adib et al., 2019). Therefore, the acceleration of improvement in improving the quality of course institutions must be carried out by evaluating the program being implemented in-depth.

Generally, evaluation is the comparison between the goal of the program and objectives that have been achieved in the program (Topno, 2012). Thus, the goal is the important component that should be appropriately planned. Different from the previous one, according to Sahayu and Friyanto (2019) believe that solving a problem that occurs in the program is not only enough by designing goals and needs analysis; it is also necessary to see the level of satisfaction of the course participants. Course participant satisfaction is the most important aspect that will influence the motivation and success rate of the program (Dewi & Kartowagiran, 2018). Correspondingly, Choudhury and Sharma (2019) state that in implementing an evaluation at the course and training institution, it is necessary to see the program's effectiveness and benefits for the company and participants. The benefits felt by course participants can be seen through the participants' goal achievement and the satisfaction of services provided by the course institution. Thus, to have good training and courses program, the goals of the program and the participants' marking about the program should be prepared. To broadly and deeply see an institution's management, such as goals, service satisfaction, and outcome of the program, the appropriate evaluation model that can be used is CIRO (Context, Input, Reaction,

and Outcome) evaluation model. The CIRO evaluation model of course and training emphasizes the problem domain and program performances as a step to design the objective, improve performance, and gain recognition (Sutton, 2006). Thus, the CIRO evaluation model is suitable to be used in this research, so it underlines the importance of the study on evaluating the management of TOEFL preparation course at one of the English village courses in East Java. The components of the evaluation model are context, input, reaction, and outcome of the program. Those components are chosen based on the importance of the institution's preparation and participants' response to the implementation of the TOEFL preparation program.

## RESEARCH METHOD

This study used qualitative and quantitative research methods. These methods were chosen because this study used the CIRO evaluation model, which required detailed data and measurement percentages. A qualitative method was used to examine the context and input aspect, while a quantitative method was used to measure the percentage of the course participants' reactions and calculate the results of their achievements. This research was conducted from February to June 2020 in the English village of Pare, Kediri, East Java.

The subject of this study was the program director, four tutors, and 30 course participants at the famous course in Pare English village. The informants were selected using the purposive sampling technique, which allowed researchers to get detailed information. The data collecting technique in this research were observation, interviews, questionnaires, and documentation. In the qualitative research, the data collected from observation and interviews were validated by triangulation. In the quantitative research, the questionnaire was validated by content validity utilizing Gregori formula by the result of 0.96. The construct validity was analyzed using SPSS utilizing the *Kaiser Mayer Olkin Measure of Sampling Adequacy (KMO)* by the result of 0.97, followed by the reliability testing analyzed using SPSS employing Alpha Cronbach's formula by the result of 0.70. It proved that the instrument is proper to use in this research.

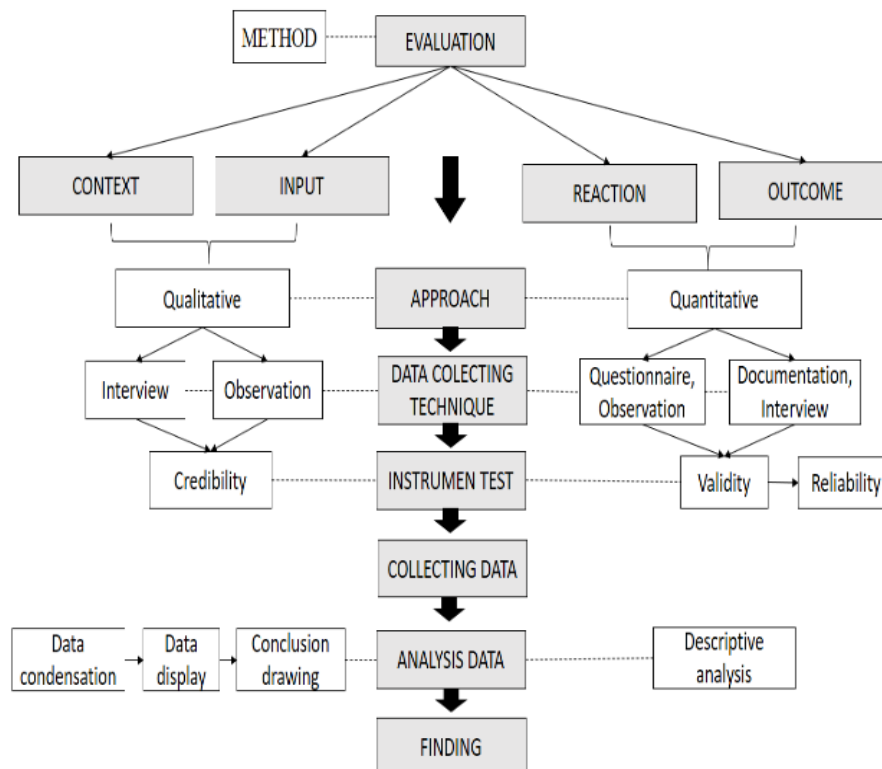


Figure 1. Method Schemes

Data analysis techniques were divided into each method used. Qualitative methods were analyzed through three stages, namely data condensation, data display, and drawing conclusions. In addition, the quantitative method uses descriptive analysis techniques based on normal distribution using Ms. Excel 2016 program to determine the percentage of responses and the learning outcomes of course participants. The details of the schemes used in the study are shown in Figure 1.

## FINDINGS AND DISCUSSION

The evaluation model used in this study is the four aspects of CIRO, which have different data collection methods. Therefore, it is explained step by step based on these components, namely context, input, reaction, and outcome, elaborated as follows.

### Context

The first stage of the evaluation is the evaluation of the context, which is divided into two aspects, namely the ultimate objective and the immediate objective in the management of the TOEFL preparation program. The program's ultimate objective is the final objectives designed by the course institute to achieve progress in the implementation of the program. The decision evaluation making in the aspect of context evaluation in the TOEFL preparation program is then compared to the standard procurement of courses that apply in Indonesia.

The research results by collecting data using interviews with the institution's director and confirmed through the tutor who taught at the institution have been confirmed verbally. The objective of the TOEFL preparation courses program has been designed without a written document. It can be interpreted that there is no official document regarding the purpose of establishing the TOEFL preparation program. It can be seen by the director's statement, "The program objectives were designed a long time before the program was implemented, but they were not written or recorded about the objective of the program." The director maintains that the objective was constructed by coordinators of the program. It is confirmed to the tutor that the program's objective is a collision between the program coordinator and the research and development department at the institution.

Based on the research interview, the ultimate objective of the TOEFL preparation program according to the tutor of the program is:

*"The final goal of the program? Yes, the course has several goals, such as helping students to achieve the score they want, it is like to make students pass in real test because the final result is not here but the real test".*

Additionally, the program director and student both argued, "The ultimate objective of the TOEFL preparation program is to pass the TOEFL real test to get a certificate which can be used for looking for job or college approval requirements". Based on the statement, the ultimate objective of the TOEFL preparation program is to help the participants pass the real test and get the score they expect.

The TOEFL minimum score of the course is 450, considering the average standard of Indonesian universities and job requirement. However, this is not an absolute value since many participants have different targets according to their needs. Another ultimate objective of the TOEFL preparation program is to help participants answer similar questions even though the questions are not the same. Based on an interview with a tutor, the TOEFL standard is as follows. "It is based on the agreement of the student's own needs". It means that participants and institutions are interrelated and support each other in increasing students' achievement.

TOEFL preparation program at these institutions is suitable to the [Regulation of the Minister of National Education No. 49 of 2007](#) concerning educational management standard in nonformal education units, in which such formulation of objectives must be planned before

the program is started. Discussions about program procurement objectives and program planning are better held regularly, once or twice a year (Gonçalves & Chauma, 2020). Such activities have been done to improve the quality of the program implementation. The final objective setting has been carried out to determine the goals that will be achieved after implementing the program, even though it is not written in an official document. Correspondingly, the lack of documents in determining the objectives is similar to the English language program, such as the AADU program, which does not have clear documents or program objectives; this has resulted in a lack of communication between stakeholders (Aktaş & Gündoğdu, 2020). The result of the research states that there is a good communication between the coordinator and program implementer about program planning, but the preparation of the documents is one of the crucial aspects in program planning, which aims to anticipate problems.

TOEFL planning program of English village course institution is adapted to the course participants needed as having a TOEFL score that can be used according to the personal target. It can be concluded that there is no score determination standard in improving learning because of the differences in the needs and basic abilities of the course participants. Similar to this study, Zhao (2020) states that attention to the needs of the course participant is the most important thing, which is useful for helping the course participants achieve the targeted competencies (Syakur et al., 2020). Further, Liu (2020) elaborates that omitting the needs of participants can affect the quality of learning and participants' satisfaction with the program. Emphasizing the course participants' needs can be considered the right step to be implemented to form the final objective so that program planning has been implemented properly. It has been implemented in the courses that have been researched.

The immediate objectives are skills that the courses provide to the participants to support the final objective. Based on the interview, the special skills students want to achieve in the TOEFL preparation course are structure skills. However, other interview results show that the participants are not all weak in this structure field. The program director states:

*“The specific objective leads to the goal of how members can reach the TOEFL score according to the required standards, but the structure is more difficult from the others, so two meetings are given a day, and listening and reading are given one time in one day”.*

As seen in the excerpt, structure skills have special attention due to the amount of material that must be discussed more because of the difficulties and demands of the participants. It was confirmed to participants of the TOEFL program who attended the courses. From the whole of the result, it can be concluded that the institution's immediate objectives are giving the skills to improve the participants' scores in TOEFL. The skills are given, including listening, reading, and structure. Based on the participants' needs, the structure class is held two classes per day while listening and reading are held one class for each per day. The meeting frequencies are four times of discussion or theory in one week, then one scoring a week, so that the total meetings in the course become 68 meetings in one-course period.

The interview results conclude that research and review have been carried out prior to the design of program objectives. The skills presented in the course are planned based on the needs of the community who would attend the course. This is parallel to the theory that paying attention to skills according to the needs of participants is important in improving the learning outcomes of the course participants (Silva & Tosqui-Lucks, 2020). As such, the steps that have been carried out by the institution are the right steps in increasing the achievement of the course participants' scores.

One of the ways is that the institution provides intensive classes that have four meetings in one day. It would give effective class, but intensive classes risk fatigue in learning course participants, so that appropriate strategies and arrangements are needed in the learning process. According to Wenjie (2020), organizing learning slots every day will be a challenge faced in order to avoid physical, mental fatigue that can interfere with the learning effectiveness. The

dividing of the learning schedule is something that becomes the attention of the program. In these courses, the timing of the TOEFL program has been declared in a good way, but it has to be improved in regulating class distances because of the long distances between each class so that it will give the teacher enough preparation time.

## Inputs

Inputs are needed to support program implementation. Common inputs that are important to maintain are school resources, teacher quality, facilities, and learning outcomes (Britton & Vignoles, 2020). Input is an important factor in the management of a course institution. The results of the interview show that the preparation of the institution towards implementing learning activities is as follows.

First, creating a program plan: the program team prepared a program planning by referring to the students' needed and TOEFL real test. Program planning was written on the syllabus, which was outlined in the learning module. However, the teacher should shape the class conditions in the learning process. The syllabus was used to synchronize the methods and strategies the tutor would use in class. Several aspects were covered in the program plan at the institution, such as learning schedule, learning model, learning assessment, and learning regulation. The examples of data obtained through analysis and documents checking are as follows.

*"Program planning refers to a learning method that is possible to use in the class that can be controlled by the teacher in the class. Then the strategies were prepared by the teacher before. All things have been exhaustive in the program or lesson planning. The mentor allows the participants to follow the syllabus by seeing the class condition. The tutors can improve the learning strategies based on the possibilities."*

Based on the interview, it can be seen that the learning plan has been determined, but it has to follow and adapt to the actual conditions in the classroom. Furthermore, all tutors and program director have the same opinion about the schedule of the program that:

*"For the admission program, those were 10 and 25 periods, so period 10 is the entry period approaching the 10th and 25th period which approaches the 25th, it is not always 10th, but sometimes we will start in the date that is approaching the 10th. For example, the 10th is on Saturday, so we will start on Monday, or if the 10th is on Tuesday, we start on Monday also".*

Learning program, one of the aspects that are considered to start learning, provides the determination of the schedule and period parallel to the learning period agreed upon in the English village. The learning model, learning assessment, and learning regulation have been well prepared and are suitable for the procurement of the course programs in Indonesia. The interview and document about the aspects of the planning program can be summed up in Table 1.

Table 1. Contents of the Planning Program

Aspect	Planning
Learning schedule	Classes would start on the 10th and 25th of every month.
Learning model	The learning model has been based on the applicable operational standards but can be modified according to class conditions.
Learning assessment	The assessment was given to participants, and evaluation was given to the tutors. Participants took the test once a week on Friday, called Scoring. Evaluation for tutors is held every week and every month during the briefing process.
Learning regulations	There were two learning regulations given in the TOEFL program: rules for tutors and the course participants. The rules for participants were the rules for the course environment that must be obeyed by course participants. Meanwhile, the regulation for tutor emphasizes the activities, duties, and responsibilities of the tutor in the learning process.

The result of this study is similar to previous research, which states that learning planning must match the students' needs and the syllabus (Perez & Mardapi, 2015). To avoid misunderstanding and misuse of instructions, themes and lesson plans are required to be formed by the program organizer (Liu, 2020). It can be concluded that the institution has designed the planning program well by looking at the aspects included in the learning plan and those data.

Second, tutor qualification: teaching tutors must match the qualifications given by the institution or national regulation. Based on the interview results, two aspects are considered in the preparation of tutors to teach in the TOEFL class: the qualifications of tutor candidates and the development of tutor candidates' skills. In the TOEFL program, the educational background is not the main problem of being a tutor. The qualification of tutors who teach in the TOEFL program is to have a minimum score of 550 in the TOEFL real test, as cited in the director program's statement:

*"The quality standard for educators in the TOEFL program is to have TOEFL score (the TOEFL standard of above 550). Some tutors do the Real Test within a minimum score of 550; then they can directly teach in the class (approximately 60% tutors in the institution have taken the real test)".*

Despite the tutor saying something conversely, the tutor revealed that his TOEFL score was 500 the first time he joined the institution. Tutors who had a score of 500 were allowed to teach, but they had to attend a similar course and were willing to participate in skills development. The tutor put in a statement, "The teachers must have studied TOEFL before and got a TOEFL test with a minimum score standard of 500". It can be concluded that tutors can be accepted in the institution even though the initial test results are insufficient because there is a development skills program that the tutor must master before starting to teach. Development is a process to improve the quality of tutors in teaching; it is called Briefings. "... it is called Briefing, which is a kind of training to improve the tutors' mastered skills". Correspondingly, another tutor explained that:

*"General development was holding through briefing, tutors in basic skill are developed to be able to teach at the second level such intermediate levels, tutors in intermediate levels are developed to the upper intermediate or to the advanced level, they are developed slowly and gradually, although when they become seniors, they will be developed more until they become briefers or trainer."*

It is concluded that the skills development process at the institution for teachers is that basic tutors would have their teaching quality improved to an intermediate level, then they would be upgraded to upper mediate or advanced. Skills development efforts at the institution are also carried out by conducting comprehensive training every year.

Third, preparation for admission of course participants: the criteria for students allowed to take the TOEFL course have been found by interviewing the director of the TOEFL program. Based on the research results at this institution, the institution did not limit course participants to a certain age or ability level. The program director said:

*"We accept whoever wants to join the TOEFL program then we would hold a pre-test before the class is started so that the classes are divided according to the results of the pre-test. For example, participants who got 400 score test and below of are treated in one class: (they will get more theory or basic English will be deepened), participants who got 400-450 score is gathered into one class, then participants who got 450 score and above will be in one class".*

It can be concluded that there are no specific criteria that participants must fulfill in order to register as prospective course participants. Nevertheless, it does not mean that the institution is unaware of the circumstances and participants' abilities in class. To reduce gaps in learning of TOEFL preparation class, the institutions conduct a pre-test prior to the class. The aim of the test is to group the course participants according to the basic abilities that they have had before taking the course.

There are three class categories provided in the implementation of the TOEFL preparation class: firstly, a class devoted to participants who have a test score of less than 400; secondly, the class for participants who already have basics and have a test score of 400-450; thirdly, the class for participants who already have scored above average skills, such having a minimum of 450 scores on the pre-test. Each class receives special treatment, methods, and material according to their basic skills.

Another aspect considered in preparation for admission of the course participants by the course provider is transparency and suitability of admission of the course participants to the infrastructure provided in the course. The implementation of the pre-test is carried out by transparency of scores, thereby creating trust between course participants and course providers is essential. Admission of prospective course participants is not given specific qualifications, but it is adjusted to the facilities available at the institution. The facilities that have been given include classrooms that are comfortable for use. The course has provided 20 classrooms filled by 15-20 participants in one class—supported by available instructors such as 37 teachers each month. The program implemented at the time of the study has three classes provided, with each of them having ten participants per class. This clearly shows that the organizer in the institution has done good preparations.

Fourth, the preparation of facilities that would be used in learning: infrastructure is evaluated by looking directly at the quality and quantity of the infrastructure available at the institution. The planning and control of infrastructure facilities are carried out by the operational division and the program team. Meanwhile, the evaluation is carried out by the management team supported by another team. Evaluation has been done every week and every month for seeing the things that need improvement quickly. The director of the program argued that "Maintenance of the infrastructure is carried out by the management team, namely the program director, operational director, and marketing director. It is the operational division which has arranged the existing facilities professionally." The statement suggests that every director in the program has their respective roles in managing existing facilities. In maintaining the available facilities, socialization is given to tutors as the users of class facilities. Socialization is held once a month. Then, the socialization of the use of the facilities should carry forward to course participants; it is used to protect the institution's facilities.

The classroom and learning environment infrastructure are as follows. One classroom has been provided for a maximum capacity of 20 people, which is equipped with a noise reducer that supports listening learning. The institution provided 20 active speakers that are used in class for listening lessons. Based on the interview results, the sound system's quality is good, while the facilities available in the classroom have been assessed based on the standard of infrastructure for managing the courses as written in the [Regulation of the Minister of National Education No. 49 of 2007](#), presented in Table 2.

Table 2. Infrastructure Quantity and Quality

Quantity		Quality	
Total Score	54	Total Score	47
Maximum Score	75	Maximum Score	75
Percentage	72%	Percentage	62%

Based on these results, the percentage of the quantity of the course infrastructure is declared good enough and meets the applicable standards. The quality of the facilities and infrastructure is not the same as the result of the quantity calculation. However, the quality of the facilities can be said in a good line.

It can be concluded that the input provided by the course has been well prepared and complies with the national government standard of non-formal education management. Thus, [Sahayu and Friyanto \(2019\)](#) state that the classroom arrangement, tutor proficiency training,



the participant needs analysis, and material designs are the things that must be given more attention in managing an institution. Fulfillment of the participants' expectations and needs can be done by dividing participants into classes according to their abilities, giving a class for students of less than 30 course participants, and setting flexible schedules that allow participants to choose according to their needs (Wenjie, 2020). In the course management, the supporting components of the implementation of training and courses must be adjusted and well prepared as the situation that institutions have carried out in this research. This is carried out in accordance with applicable standards, and it is well accomplished by regular maintenance of the facilities.

## Reaction

The assessment of participant reactions is divided into two aspects: participant satisfaction reactions to the TOEFL program and participant assessment reactions to the implementation of TOEFL learning in class. First, participants were allowed to fill out participant satisfaction questionnaires to assess the infrastructure provided, participant satisfaction reactions to the tutoring service, participant satisfaction reactions to the material presented, participant satisfaction reactions to the methods used in class, participant satisfaction reactions to the time management provided, and participant satisfaction reactions to training activities of the TOEFL questions implemented. The total number of questions on the satisfaction questionnaire is 36 statement items. The result of the interview to the participant will confirm the data; the number of questions is six questions. The results of the questionnaire obtained are depicted in a pie diagram in Figure 2.

**Participant Satisfaction Level  
of the TOEFL Program**

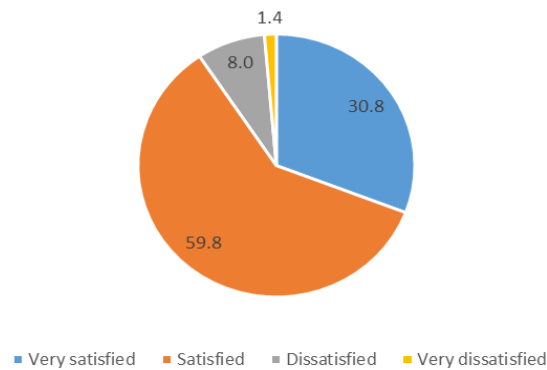


Figure 2. Participant Satisfaction Level

The data were generated through data analysis using a flowchart in the Microsoft Excel 2016. The data are the analysis result of course participants' satisfaction with the TOEFL program at a course institution that has been conducted for a month. The data show that the participants are very satisfied 30.8%, 59.8% satisfied, 8.0% dissatisfied, and 1.4% very dissatisfied. Thus, most of the course participants are satisfied with the implementation of the TOEFL program in English Village.

Based on the participants' interview, they are satisfied with the infrastructure provided because the class is comfortable, the facilities are adequate, modules have been fulfilled, media and teaching materials have been adjusted. In contrast, the participants are not satisfied with the existing facilities because it is sometimes hot in class during the day. It could be seen based on participant's statement, "Classroom conditions are quiet, comfortable, and have adequate

facilities”. Participants are satisfied with the tutor because the tutor provides the opportunity to ask questions for the participants, but one of the tutors sometimes gave the material fastly. Satisfaction with the material in the TOEFL program is caused by the material that is appropriate to the students’ level, even though some of them feel that the material is difficult. One of the participants stated that, ”The tutor teaches well using a good method, but sometimes the material is too hard for me”. The method used in the material teaching makes students understand it well. The schedule is prepared well based on the students’ needs, but they are disappointed because they should stop the program before one month because of the Covid-19. The participants are satisfied with the practice activities held in class since the tutor gives the tips and tricks to answer the TOEFL test.

Second, participants’ reaction in the implementation of learning is taken to see the participants’ assessment of the activities carried out in class. Assessment of the learning process is differentiated according to the class that has been carried out, which has different assessment results. Four classes are assessed: listening class, reading class, structure 1 class, and structure 2 class. The number of questions that course participants must answer is 17 questions in each session. The score for the course participant’s assessment of the TOEFL learning process is categorized based on Table 3.

Table 3. Assessment Score Category

Category	Interval
Very good	$X \geq 54.74$
Good	$54.74 > X \geq 42.50$
Bad	$42.50 > X > 30.26$
Very bad	$X < 30.26$

Based on calculations using Microsoft Excel, the score of the course participants’ assessment of the learning process is 55.56. It indicates that the results of data analysis of the four subjects that the participants have followed have an average score of 55.56 in the very good category. The research findings support this during observation; the result of the observation value is 9.70, with a very good category. Thus, it can be concluded that the learning process has gone very well.

It can be summed up that participants of the TOEFL preparation course program are included in the satisfactory category. Most of the participants are satisfied with the services provided by the institution. All learning sessions are in good categories. These results are different from previous research on the evaluation of the English for a specific exposure program. Participants gave a negative view of learning that did not improve the participants’ skills. The problems found were generalized classes, inaccurate methods, inappropriate materials, and also inappropriate syllabus designs to the needs of the participants (Alemi & Pazooki, 2020).

It can be said that the reaction results depend on the responses and ratings that are given by the participants. Thus, to avoid future problems, quality improvement must be continued. In line with this, Nazri et al. (2020) believe that participants must have good strategies and techniques in answering the TOEFL questions. However, most of the problems that arise are that teachers provide strategies to answer questions but they ignore the atmosphere and comfort built in the classroom. Further, Barnes (2016) explains that the TOEFL preparation class is very structured and goal-oriented. Therefore, teachers often limit the teaching styles and methods that are used. The results of this study show that the data are different from the aforementioned previous studies. The teachers in these institutions have taught well, and they can adjust the class to the condition so that the course participants do not get bored in learning. This can be reflected in the results of the learning process questionnaire, which achieve very good grades.

## Outcome

Scoring is an assessment carried out once a week. Scoring is done to see the learning progress of the course participants. There are three skills included in the TOEFL test: listening, structure, and reading. These three skills are combined in one TOEFL test. The TOEFL test results that are calculated are the pre-test, scoring 1, scoring 2, and scoring 3, as described in Table 4.

Table 4. The TOEFL Test Results

Time	Comparison Score	Average Value	Score Increasing
Pre-Test	365	40	Initial score
Scoring 1	378	44	Increase
Scoring 2	391	48	Increase
Scoring 3	372	42	Decreased

Based on Table 4, the pre-test score or initial test was 365 with the correct number of 40 items. In the first scoring, it increased to 378 with the correct number of 44 items. The second scoring increased from the previous average score to 391 with 48 correct points, and the third scoring decreased to 372 with 12 correct points. Thus, there was a decrease in the score at the end of the course period.

The results of interview with course participants state that the improvement in learning is most felt in learning structures, even though some course participants felt that their TOEFL scores increased more in listening and reading skills. Some participants state they have an increase in learning even though the scores they get do not increase at the time of the TOEFL test. The problems faced during the exam are the speaker voice that is not heard, the processing time is too fast, and some participants come late, which disturb concentration.

Several problems can occur during the test, such as the results of this study. According to [Gür and Eriçok \(2020\)](#), understanding the material's content is the problem that can inhibit increasing the TOEFL test score. However, the institution in this study has anticipated this problem by dividing the participants based on the basic skills seen in the pre-test before the TOEFL class was held. The problems also occur during the test. [Yogawati and Widihastuti \(2019\)](#) state that students often face problems during exams, such as nervousness, low motivation, pessimism, mastery of vocabulary, low self-confidence. Another theory problem that often arises in exams is anxiety and stress fatigue, which can make it difficult for participants to answer questions ([Nikolaieva, 2016](#)). Problems often occur during the TOEFL test implementation, so it must be considered to reduce problems that will arise. Generally, it can be stated that the planning and implementation of the TOEFL test preparation program have been carried out properly by the institution.

## CONCLUSION

Based on the research findings and discussion, the evaluation of the TOEFL program in the English village can be concluded as follows. The evaluation context is divided into the final objective and the program's ultimate objective, which are the ultimate objective. The context of the program has been fulfilled in accordance with the applicable national and real test standards. The input of the evaluation is well prepared, including a program plan that provides some aspects. Tutor qualification improvement is carried out every week and month, which is called the Briefing program. The admission of the course participants is based on a pre-test to see the participants' initial abilities. The facilities of the TOEFL program are in a good result; the quantity assessment score is 72%, and the quality is 62%. It can be said that the infrastructure has met the existing standards. However, it needs to be further improved and repaired in terms of its quality and quantity.

The reaction of participant satisfaction in the satisfied category is 90.6% and is 9.4% dissatisfied. It states that the TOEFL program is irresistible because many course participants are satisfied. The reaction of the participants to the learning process is in the good and very good categories. Generally, learning is carried out according to the learning procedures, so it is declared that the learning goes well. Learning outcomes increased in the first and second scoring but decreased in the third scoring with the same results as the pre-test. However, the decision in this evaluation is seen as a whole. It is concluded that the institution has implemented the program in accordance with the applicable standards.

## REFERENCES

- Adib, H. S., Mardapi, D., Zamroni, Z., & Jait, A. (2019). Evaluation of Islam education teachers training implementation. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 23(2), 106–116. <https://doi.org/10.21831/pep.v23i2.20986>
- Aktaş, C. K., & Gündoğdu, K. (2020). An extensive evaluation study of the English preparatory curriculum of a foreign language school. *Pegem Eğitim ve Öğretim Dergisi*, 10(1), 169–214. <https://doi.org/10.14527/pegegog.2020.007>
- Alemi, M., & Pazooki, S. J. (2020). *A stakeholder-based evaluation of engineering ESP courses' effectiveness*. <https://doi.org/10.21203/rs.3.rs-58769/v1>
- Ali, H. (2012). The use of silent reading in improving students' reading comprehension and their achievement in TOEFL score at a private English course. *International Journal of Basic and Applied Science*, 1(1), 47–52. <https://doi.org/10.17142/ijbas-2012.1.1.8>
- Barnes, M. (2016). The washback of the TOEFL iBT in Vietnam. *Australian Journal of Teacher Education*, 41(7), 158–174. <https://doi.org/10.14221/ajte.2016v41n7.10>
- Britton, J., & Vignoles, A. (2020). Education production functions. In G. Johnes, J. Johnes, T. Agasisti, & L. López-Torres (Eds.), *Handbook of contemporary education economics* (pp. 246–271). Edward Elgar Publishing. <https://doi.org/10.4337/9781785369070.00016>
- Choudhury, G. B., & Sharma, V. (2019). Review and comparison of various training effectiveness evaluation models for R & D organization performance. *PM World Journal*, 8(2), 1–13.
- Dahuri, D. (2019, December 12). Indeks kemampuan bahasa Inggris orang Indonesia nomor 61. *Media Indonesia*. <https://mediaindonesia.com/humaniora/277217/indeks-kemampuan-bahasa-inggris-orang-indonesia-nomor-61>
- Dewi, L. R., & Kartowagiran, B. (2018). An evaluation of internship program by using Kirkpatrick evaluation model. *Research and Evaluation in Education*, 4(2), 155–163. <https://doi.org/10.21831/reid.v4i2.22495>
- Direktorat Pembinaan Kursus dan Pelatihan. (2018). *Data dan informasi kursus dan pelatihan 2018*. Direktorat Jenderal Pendidikan Anak Usia Dini dan Pendidikan Masyarakat.
- English First. (2020). *The world's largest ranking of countries and regions by English skills*. English First - English Proficiency Index. <https://www.ef.co.id/epi/>
- Gonçalves, G. J., & Chauma, A. M. (2020). Challenges for foreign English language teacher education programme in Mozambique with focus on Zambézia Teacher Training Colleges. *London Journal of Research in Humanities and Social Sciences*, 20(1), 23–34. <https://research.journalspress.com/index.php/socialscience/article/view/585>

- Gür, R., & Eriçok, B. (2020). The relationship among academic success scores of graded foreign language courses. *Dil ve Dilbilimi Çalışmaları Dergisi*, 16(2), 809–821. <https://doi.org/10.17263/jlls.759309>
- Ismail, I., & Othman, R. (2020). A review of literature on the English language entry requirement for international students into postgraduate programs in Universiti Teknologi Malaysia. *Journal of Critical Reviews*, 7(11), 543–549. <http://www.jcreview.com/?mno=117854>
- Kurniawati, E. (2017). Manajemen strategik lembaga pendidikan Islam dalam meningkatkan mutu pendidikan: Studi kasus di Madrasah Aliyah Nahdlatul Ulama Gondang Sragen. *At-Taqaddum*, 9(1), 113–132. <https://doi.org/10.21580/at.v9i1.1784>
- Liu, F. (2020). Failure of humanities-based instruction to achieve students' language goal in college English courses. *Journal of Language Testing & Assessment*, 3(1), 1–4. <https://doi.org/10.23977/langta.2020.030101>
- Ma, J., & Cheng, L. (2016). Chinese students' perceptions of the value of test preparation courses for the TOEFL iBT: Merit, worth, and significance. *TESL Canada Journal*, 33(1), 58–79. <https://doi.org/10.18806/tesl.v33i1.1227>
- Mardapi, D. (2017). *Pengukuran, penilaian, dan evaluasi pendidikan* (Revised ed). Parama Publishing.
- Maruf, Z., Rahmawati, A. S., Siswantara, E., & Murwantono, D. (2020). Long walk to quality improvement: Investigating factors causing low English proficiency among Indonesian EFL students. *International Journal of Scientific & Technology Research*, 9(3), 7260–7265. <http://www.ijstr.org/final-print/mar2020/Long-Walk-To-Quality-Improvement-Investigating-Factors-Causing-Low-English-Proficiency-Among-Indonesian-Efl-Students.pdf>
- Mulyasa, E. (2005). *Menjadi guru profesional*. Remaja Rosda Karya.
- Nazri, M. A., Wijaya, H., & Zainurrahman, Z. (2020). EFL students' ability in answering TOEFL reading comprehension section. *Journal of Physics: Conference Series*, 1539, 012044. <https://doi.org/10.1088/1742-6596/1539/1/012044>
- Nikolaieva, O. (2016). *A qualitative study on preparing EFL students to take the TOEFL internet-based (iBT) test in the Ukrainian context* [Master thesis, University of Stavanger]. <https://uis.brage.unit.no/uis-xmlui/handle/11250/2400189>
- Nurhayati, N., Hendrawaty, N., & Angkarini, T. (2013). The acquisition of English as a foreign language in Pare East Java (Kampung Inggris) (A case study of what and how the acquisition of English in Pare). *Deiksis*, 5(2), 81–88. <https://journal.lppmunindra.ac.id/index.php/Deiksis/article/view/462>
- Perez, B. E. O., & Mardapi, D. (2015). Evaluation of the bridging course offered at a university to foreign students: Batches of 2012 and 2013. *Research and Evaluation in Education*, 1(2), 146–157. <https://doi.org/10.21831/reid.v1i2.6667>
- Regulation of the Minister of National Education No. 49 of 2007 on the Standard of Educational Management by Nonformal Educational Units, (2007).
- Sahayu, W., & Friyanto, F. (2019). The effect of Youtube on high school students' second language acquisition. *International Journal of Linguistics, Literature and Translation (IJLLT)*, 2(6), 38–44. <https://doi.org/10.32996/ijllt.2019.2.6.5>

- Sakurai, N. (2020). Exploring a placement test for extensive reading programs. *Humanities Series*, 2(3), 53–70. [https://ksu.repo.nii.ac.jp/?action=repository\\_action\\_common\\_download&item\\_id=10468&item\\_no=1&attribute\\_id=22&file\\_no=1](https://ksu.repo.nii.ac.jp/?action=repository_action_common_download&item_id=10468&item_no=1&attribute_id=22&file_no=1)
- Silva, A. L. B. de C. e, & Tosqui-Lucks, P. (2020). Around the world in Aeronautical and Aviation English courses. *Revista CBTeLE*, 2(1), 418–440. <https://revista.cbtecle.com.br/index.php/CBTeLE/article/view/274>
- Sudjana, S. (2004). *Pendidikan nonformal*. Falah Production.
- Sutton, B. (2006). *Adopting a holistic approach to the valuation of learning programmes deployed in corporate environments* [Thesis, Middlesex University, London]. <https://eprints.mdx.ac.uk/2667/>
- Syakur, A., Zainuddin, H. ., & Hasan, M. A. (2020). Needs analysis English for specific purposes (ESP) for vocational pharmacy students. *Budapest International Research and Critics in Linguistics and Education (BirLE) Journal*, 3(2), 724–733. <https://doi.org/10.33258/birle.v3i2.901>
- Syamsuddin, I., & Min, A. (2014). Assessing moodle as learning management system platform for English course based TOEFL. *International Journal of Computer Trends and Technology*, 18(6), 276–279. <https://doi.org/10.14445/22312803/IJCTT-V18P158>
- Topno, H. (2012). Evaluation of training and development: An analysis of various models. *IOSR Journal of Business and Management*, 5(2), 16–22. <https://doi.org/10.9790/487X-0521622>
- Wang, Y. (Tina). (2019). The impact of TOEFL on instructors' course content and teaching methods. *The Electronic Journal for English as a Second Language*, 23(3), 1–18.
- Wenjie, S. (2020). Evaluating an English course for master students in China: A case of business English for accounting program. *International Journal of English Language Teaching*, 7(1), 31–40. <https://doi.org/10.5430/ijelt.v7n1p31>
- Yogawati, N. D., & Widihastuti, W. (2019). Evaluating the implementation of English communication therapy (ECT): An objective structured clinical assessment (OSCA) approach. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 23(1), 87–94. <https://doi.org/10.21831/pep.v23i1.22449>
- Zhao, L. (2020). A study on the feedback of college English dynamic classified teaching effect. *Theory and Practice in Language Studies*, 10(6), 678–684. <https://doi.org/10.17507/tpls.1006.08>