

AN EVALUATION ON TOEFL WORKSHOP PROGRAM USING STAKE'S COUNTENANCE MODEL

Yupika Maryansyah¹, Syukri Hamzah², Hadiwinarto³.

¹Universitas Muhammadiyah Bengkulu,

^{2,3}Universitas Bengkulu

yupikamaryansyah@umb.ac.id , shamzah@unib.ac.id , hadiwin@unib.ac.id

Received: December 29, 2022 Revised: January 25, 2023 Accepted: February 12, 2023

ABSTRACT

The Test of English as Foreign Language (TOEFL) is getting more and more important every day for both further study and career advancement. Accordingly, many educational institutions, both state-owned and private in Indonesia carried out innovative programs to increase their graduates' TOEFL scores. Universitas Muhammadiyah Bengkulu (UMB), as an educational institution also carried out a program aiming to improve its graduates' TOEFL score. The program is called as the TOEFL workshop for UMB students. This study is an evaluative study analyzing a program administered by UMB, TOEFL workshop for UMB students. The researchers used Stake Countenance Model as a method in the current study. The results of the study show that there is a contingency between antecedent, transaction, and outcome of the program of TOEFL workshop for UMB students.

Keywords: Evaluation, TOEFL Workshop Program, Stake's Countenance Model.

INTRODUCTION

The importance of TOEFL is increasing day by day for both career advancement and further studies. As a result, numerous educational institutions in Indonesia, including state-owned and private ones, are introducing innovative programs to boost their graduates' TOEFL scores. Muhammadiyah University of Bengkulu (UMB), a privately owned educational institution, is also taking part in this effort by conducting a program known as the TOEFL Workshop for UMB students, with the goal of enhancing their English language skills. To ensure the success of this program, the Rector of UMB issued a decree in 2016, which stated that all UMB students must participate in the TOEFL Workshop program.

This program has been being implemented at UMB since 2016 and is still running until now. Even though it has been held for about 6 years, not a single evaluation has been conducted on this program. There are two urgencies to carry out a comprehensive evaluation

study on the programs mentioned above. First, since it was implemented in 2016, this program has never been evaluated. Second, in an observation made by the authors, there was stagnation in increasing the TOEFL scores of UMB students. The increase is still ongoing, but at a very sluggish pace, rather than having not increased at all. Due to the issues above, it is very important to evaluate the TOEFL Workshop program for UMB students.

The results of this study are very important to be used as input for UMB related to the implementation of the program being held, in this case the TOEFL Workshop program for UMB students. Choudhury and Sharma (2019) emphasize the significance of conducting assessments of education or projects to guarantee efficiency and advantages for a group or individuals, which aligns with the given statement. After a firm evaluation result is identified, the university can consider further action on the TOEFL Workshop program for UMB students. Specifically, this study aims to evaluate implementation of UMB student TOEFL Workshop program which is held annually.

A good TOEFL score is not just a certificate, but more like a need after completing education, especially after the undergraduate level. At regional and international levels, TOEFL has become a standardized English proficiency test tool (Ismail & Othman, 2020). The TOEFL score obtained by someone can be used anywhere, both at national and international levels. Furthermore, Iftanti and Maunah (2012) argue that TOEFL is the highest achievement in learning English for a student. This means that if an undergraduate graduate gets a high score after taking the TOEFL test, then he or she will be considered proficient in English. Here we can see that the TOEFL score can be said to be a measure of someone's mastery of English (Esfandiari et.al, 2018). For foreign speakers, it is clear that they need certain training to get a high score on the TOEFL test. One of the ways to increase TOEFL score is to join a TOEFL test preparation training program in all forms; workshops, training, courses, and others (Wang, 2019). Test takers must have a good strategy in order to be able to answer TOEFL questions (Nazri et. Al, 2020).

The TOEFL workshop organized by UMB is a short training aimed at providing knowledge on TOEFL test and how to answer the test questions. The ultimate goal of this program is to increase TOEFL score which involves three parts of the TOEFL test; listening comprehension, structure and written expression, and reading comprehension (Sakurai, 2020).

As is widely known, TOEFL scores are now required for career advancement and continuing studies. It is in this context that UMB feels the need to organize programs that can accommodate the needs as described above when alumni finish studying at UMB.

The implementation of TOEFL workshop at UMB was held by an institution at UMB called as Language Center of University of Muhammadiyah Bengkulu. This delegation of tasks is based on the Rector's decision as mentioned earlier. This institution must carry out TOEFL training for all UMB students. To start the TOEFL workshop, student registration takes place after notifying all the study program heads at UMB, as part of its implementation process. It consisted of registration time, how to register, administrative requirements, and implementation time. The UMB Language Center sets the latest registration time for students, usually one month. After registration time, within one week of the registration deadline, the UMB Language Center makes preparations for implementation, such as preparing facilities and infrastructure, instructors, materials, class distribution, implementation rules, and others. Then the day before the implementation, important matters were announced to students participating in the workshop. During the workshop, the participants who are UMB students take part in 3 sessions consisting of listening comprehension session, structure and written expression session, and reading comprehension session. In each session the participants will be given material on how to answer TOEFL questions and exercises. In terms of teaching, UMB Language Center has established a TOEFL lesson plan guideline which includes; 1) main material; 2) time allocation; 3) learning objectives; 4) basic competence and indicators of achieving TOEFL competence; 5) learning materials; 6) learning methods; 7) learning resources; 8) learning steps; and 9) assessment of learning outcomes/evaluations.

Wood (2001) defines evaluation as the process of collecting and analyzing data to assess the extent of achievement of program objectives, which is then used to make informed decisions. Evaluation models can serve as a guide for management, data collection, and analysis. The evaluation of learning involves assessing the planning, implementation, and assessment processes, as well as their impact on students. Its objective is to identify and correct any shortcomings in the learning process, and it can serve as a foundation for future improvements (Mardapi, 2017; Adib et al., 2019). As previously stated, the TOEFL workshop

program for UMB students has never been evaluated, therefore, this study evaluates the TOEFL workshop program for UMB students with Stake's Countenance Model.

The countenance stake model comprises two matrices: the description matrix and the consideration matrix. The evaluator can proceed with the consideration matrix once the description matrix has been completed. The *description* matrix consists of *intent* and *observation* categories. The *consideration* matrix consists of *standard* and *judgment* categories. In each category there are three focuses: (a) *antecedent* (context), namely a condition that existed before instructions related to results, (b) *transaction* (process), which is a process of activity instructions, and (c) *outcomes* (results), namely the effect of experiences, observations and work results (Stake, 1967). The *description* matrix relates to the categories of *intent* or objectives planned in the development of a program. The observation category relates to what actually happens in the program as the desired implementation of the *intent* or purpose. The *standard* categories in the *consideration* matrix are criteria that must be met by a program that is the focus of evaluation. In the *consideration* category, the evaluator considers what has been evaluated in the *description* matrix. The entire matrix supporting the Stake model consists of 12 squares.

According to Wood (2001), Stake developed an evaluation framework to aid evaluators in gathering, structuring, and interpreting both quantitative and qualitative data. The fundamental aspect of evaluation activities is the production of information that can be used as an alternative for decision-making (Stufflebeam, 1973; Hamm, 1985; Stake, 1967; Stufflebeam & Shinkfield, 1984). The pertinent stages of Stake's evaluation framework are *Antecedent* (Input), *Transaction* (Process), and *Outcomes* (Products) (Provus, 1969; Kaufman & Thomas, 1980). Outcome description of the Stake model is the impact of implementing a program.

The countenance stake model's strength lies in its ability to distinguish between program actions and decisions during the *Input*, *Process*, and *Outcomes* stages. This model's advantage is that the evaluation is tailored to the program's specific needs, enabling a comprehensive description of learning programs and offering significant opportunities for gaining knowledge and theories related to the programs being evaluated.

Antecedent refers to the pre-existing condition that is related to the outcomes of learning, while *Transaction* denotes the process of the learning experience. There are two

methods for processing descriptive evaluation data in educational programs: identifying the linkages between *Antecedent*, *Transaction*, and *Outcomes*, and evaluating the congruity between the expected objectives and the observed conditions (Stake, 1977). It is important to determine if the execution of the learning plan matches the implementation of learning in the classroom or if any deviations are present.

Data is considered congruent when the anticipated objectives are actually achieved. When reviewing a single line of data matrix, evaluator can compare the cells containing the objectives and observations to identify discrepancies and confirm congruity. Additionally, it is crucial to examine the relationships or contingencies between variables. In program evaluation, where the goal is to enhance educational programs by identifying potential relationships, the evaluator's role is determined by identifying the conditions present in *Antecedent* and *Transaction* stages. The identification of these relationships is based on empirical evidence (Wood, 2001). The analysis of contingencies comprises of both logical analysis and empirical analysis.

Logical analysis is conducted to examine the relationship between the *antecedents*, *transactions*, and *outcomes* of the target boxes. The evaluator needs to determine if the *antecedents*, which have been identified in the development of the learning program, are achievable with the proposed transaction plan or if there is another transaction model that is more effective. The second form of analysis is empirical analysis, which follows the same methodology as logical analysis, but uses empirical data. Hamm (1985), states that the Stake Model is beneficial for providing broad evidence to support the success of a particular program by documenting reasonable relationships between all components.

Based on the given explanation, the *antecedents* in the UMB TOEFL workshop program refer to the phase of planning the learning process. In this instance, *antecedents* in the UMB TOEFL workshop program refer to the phase of planning the learning activities created by the UMB Language Center and TOEFL workshop instructor. The *transaction* or process pertains to the execution of the learning activities that are aligned with the established plan, and the *outcomes* or products refer to the achievements in the form of TOEFL scores.

METHODOLOGY

This study uses a descriptive method with the Countenance Stake evaluation model. The Countenance Stake evaluation model is an analysis of evaluation process that emphasizes two types of operations, namely *descriptions* and *judgments*, and distinguishes three phases in program evaluation, namely: (1) *Antecedents* in this study, namely TOEFL lesson planning; (2) *Transaction*, is the implementation of TOEFL learning; and (3) *Outcome*, is result of this program in form of TOEFL score obtained in TOEFL try out test. The *description* matrix relates to the intense TOEFL workshop program and the results of observations from this program. The *judgment* matrix relates to standards or criteria, in this case learning plans made by the program organizer, UMB Language Center. The greatest emphasis on this model is the opinion that evaluators make on a program being evaluated. The research design uses a countenance evaluation model developed by Stake as shown in Figure 1. The design was used on this study.

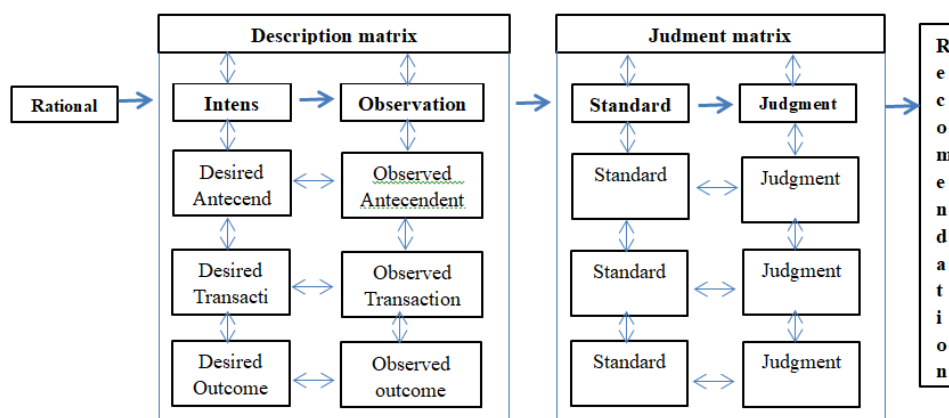


Figure 1. Countenance Stake Research Design

The Countenance Stake Evaluation model comprises four stages, starting with an initial step, followed by data collection, logical analysis, and empirical analysis. The initial step taken was to compile rationale for TOEFL workshop program at UMB. This section involves collection of preliminary data related to the TOEFL program implemented by the UMB Language Center. The data includes the planning, implementation, and TOEFL score results, as well as the theoretical foundation, regulatory support, and actual conditions that exist. The data collection stage at the intents stage is carried out to obtain information on objectives of TOEFL workshop program and expected effects of the program. The evaluation of this

objective is conducted in three parts of the evaluation components, which are the antecedent that refers to the plan that has been developed by the organizers, the transaction or implementation process, and the analysis of the TOEFL scores as the outcomes of the program. The analysis is performed by considering the objective conditions of the program and then processing the data in the description matrix using two concepts, contingency and congruence. These two concepts differ in their usage. Contingency is used to analyze data vertically, looking for relationships/congruence between antecedents, transactions, and outcomes. Contingency analysis is carried out in two ways, namely logical connectivity and empirical connectivity.

Logical analysis of data in this study is used to provide considerations regarding relationship between antecedents, transactions, and training outcomes contained in the *intents matrix*. The results of this analysis find out whether the plans made by the organizers, in this case the UMB Language Center will be reached with the proposed transaction plan. As well as regarding relationship between program implementation and the expected program results. Empirical analysis was carried out to consider relationship between antecedent, transaction, and outcome of the TOEFL score. This analysis is based on empirical data obtained in the field.

In addition of seeking for contingencies, researchers also provided considerations regarding congruence or differences that occur between what is planned and what happens in the field. To perform the congruence analysis, a set of clear and measurable criteria is developed to evaluate the program implementation at all stages of the evaluation. The standards to be used are based on theoretical considerations and practical considerations in the research field conditions. Analysis of conformity of standards with research data will be basis for decision making (judgment). This decision making was carried out for three evaluation components, antecedent, transaction, and outcome. The last step of this research is to provide recommendations and considerations based on evaluation results.

Data collection techniques in this study used observation, documentation, and interview techniques. Observation was used in the implementation of training using observation sheets in accordance with process standards. The act of observing was performed by researchers who acted as observers. This was done so that the observer's credibility can be trusted. To evaluate the lesson plans created by the organizers, documentation was utilized. The evaluation involved assigning a score based on the organizers' quantification plan, using

an assessment instrument that followed the process standard. Interview was conducted with the organizer, UMB Language Center and training instructors to obtain data on the readiness of implementing the TOEFL workshop program for UMB students. The results of the interviews serve as triangulation data supporting observation results.

Data analysis in this study was carried out in a qualitative descriptive manner. The researchers used thematic analysis to analyze the qualitative data, comparing the data from the three stages of the Stake model (antecedent, transaction, and outcomes) in the description matrix to the standards in the consideration matrix. The conclusions were then drawn based on this analysis. In the thematic analysis, an analysis flow was followed which consisted of data collection, data reduction, and verification conclusions. Data on the assessment of planning documents and the implementation of learning were analyzed in descriptive statistics using percentage formula as follows:

$$\text{Score} = \frac{\text{Raw Score}}{\text{Maximum Score}} \times 100$$

The percentage calculation value (%) is then converted in qualitative form to determine achievement actuality. The range of values is presented in Table 1.

Table 1. Conversion of program achievement calculation results

No	Percentage Scale	Value Category	Evaluation Result Category
1	86-100	A	Excellent
2	71-85	B	Good
3	57- 70	C	Sufficient
4	NA-56	D	Poor

The percentage of achievement scores shows how much UMB students' TOEFL workshop program implementation can be achieved. The percentage of the achievement score is then used to describe suitability between standard of program implementation processes that have been set with results found in the field.

RESULT AND DISCUSSION

The findings of the evaluation process for each stage are presented in a table format using the Countenance Stake matrix. The tables are grouped according to the three program components: antecedents, transactions, and outcomes. The tables contain information about intensity, observation, standards, and judgment for each of the program components. After that, the congruence and contingency analysis is performed.

Congruence

Antecedent Component

The component evaluated in the antecedent is the lesson plan made by the UMB Language Center. Below is Table 2 which shows the Countenance Stake matrix for the antecedent components.

Table 2. Countenance matrix of antecedent components

Description Matrix		Judgment Matrix	
Intent	Observation	Standard	Judgment
The lesson plan made by instructor is in accordance with the TOEFL preparation Lesson Plan Guideline made by UMB Language Center.	The actualization of the achievement of lesson plans made by TOEFL workshop instructors was 69% in the sufficient category, not all TOEFL instructors planned TOEFL lessons in accordance with the standard lesson plans made by UMB Language Center.	Components of lesson plan prepared by UMB Language Center includes; 1) main material; 2) time allocation; 3) learning objectives; 4) basic competence and indicators of achieving TOEFL competence; 5) learning materials; 6) learning methods; 7) learning resources; 8) learning steps; 9) assessment of learning outcomes/evaluations.	Some of lesson plans made by TOEFL instructors in TOEFL workshop program at UMB Language Center is not in accordance with the learning plan guidelines developed by UMB Language Center. Especially on the time allocation indicator.

Based on Table 2, the lesson plan made by instructors for the TOEFL workshop for UMB students is at the *sufficient* category level (69%). Intense agreement with observations, in the description matrix found a discrepancy in the time allocation designed by the instructors with the TOEFL lesson plan guidelines by UMB Language Center. In the end, this causes a disturbance in the evaluation of learning at the end of the session, indicating that the evaluation indicators were not met satisfactorily. In summary, there are two indicators that the UMB student TOEFL workshop instructors have not fully met. There are still instructors who have not followed the plan guidelines set by the organizers. Sometimes the time allocation provided cannot accommodate the material presented. The time allocation given by the organizers is for 2 full hours for each session taught; listening competence, structure and written experience, and reading comprehension.

Among what often happens is when the instructor has not allocated enough time to provide training activities for sessions being taught. This fact was confirmed from the results of interviews with the UMB Language Center representatives as the organizer. They stated that instructors were often short on time so practice was often skipped. This is significant considering that learning TOEFL requires a lot of practice. It can be seen that the instructor spends too much time explaining strategies for answering questions rather than giving practice time to trainees. TOEFL training class like this is not recommended because it will make the TOEFL learning class boring. On the other hand, in other indicators, subject matter, learning objectives, basic competencies and indicators of achievement of TOEFL competencies, learning materials, learning methods, learning resources, and learning steps have been running as expected by the organizers.

Transaction Component

Table 3. Countenance matrix of transaction components

Description Matrix		Judgment Matrix	
Intent	Observation	Standard	Judgment
The instructor has carried out learning process based on TOEFL preparation Learning Plan Guidelines developed by UMB Language Center.	The actual achievement of implementation of TOEFL preparation learning is 60%, which is in the <i>sufficient</i> category, not all TOEFL instructors plan TOEFL learning according to the standard lesson plan made UMB Language Center.	Implementation of TOEFL workshop learning based on TOEFL learning plan guidelines includes: (1) introduction: a) preparing workshop participants psychologically and physically to participate in the learning process; b) giving motivation to learn students contextually according to the benefits in everyday life; c) asking questions that relate previous knowledge to the material to be studied; d) explain the learning objectives or basic competencies to be achieved; and e) conveying the scope of material and explanation of activity descriptions according to the syllabus. (2) Core Activities: using learning models, learning methods, learning media, and learning resources adapted to the characteristics of workshop participants and materials. (3) Closing Activities. In the closing activities, the instructor and the workshop participants, both individually and in groups, reflect to evaluate: a) the entire series of learning activities and the results obtained so that they jointly find direct or indirect benefits from the learning outcomes that have taken place; b) provide feedback on the process and learning outcomes.	The implementation of workshop learning is not fully in accordance with the guideline standards provided. Some instructors still need to improve their professionalism through instructor training activities that will increase their professionalism in teaching TOEFL training materials.

able 3 shows that the level of achievement in learning implementation falls into the *sufficient* category with a score of 60%. The analysis revealed that there was a lack of conformity between the actual implementation of the learning program and the existing TOEFL preparation lesson plan guidelines. The discrepancy was discovered in the preliminary stage of the program, specifically in the areas of asking thought-provoking questions, communicating the benefits of the learning materials, demonstrating something related to the theme, and checking initial behavior. This causes workshop participants to sometimes be unable to relate their background knowledge to the workshop material. The result is that while participating in the workshop, the participants seem to lose attention and do not understand what is being conveyed.

The incompatibility of the implementation of TOEFL learning with standards set by instructor stems from the lesson plan they make, the instructor, learning support factors such as media, and teaching methods. The lesson plan made by instructor is not optimal, due to the limited ability of the instructor in updating learning methods and strategies, and the instructor seems to have difficulty managing time. This has an impact on the instructor's way of teaching which always only forms a culture of memorizing rather than forming the critical thinking pattern of the trainees. In fact, material that is rich and optimizes technology such as YouTube or websites on the internet will make it easier for trainees to understand. Theoretically, a situation where the method is not suitable, the material is not suitable, and the wrong learning planning design can form negative perceptions of the trainees (Alemi & Pazooki, 2020).

Outcome Component

The component evaluated in the *outcome* was the TOEFL workshop learning outcomes obtained from TOEFL score from TOEFL try out test held on the last day of the TOEFL workshop at TOEFL workshop for UMB students batch IV administered in November 2022. The table below presents countenance matrix of *outcome* components.

Table 4. Countenance matrix of outcome components

Description Matrix		Judgment Matrix	
Intent	Observation	Standard	Judgment
The results of TOEFL workshop learning is	The actual achievement of TOEFL learning outcomes is 65%, in the	High score of TOEFL, 677.	The learning outcomes of workshop participants have not met the

obtaining a maximum score of 677.	<i>sufficient</i> category. It was found that some workshop participants did not get maximum score on the TOEFL try out test.		objectives of the workshop. Participants are not fully aware that TOEFL score is greatly influenced by adequate practice frequency.
-----------------------------------	---	--	---

Table 4 explains that there is no match between learning outcomes at the TOEFL workshop based on the highest TOEFL score, namely, a score of 677 (Phillip, 2006). The achievement of students' learning outcomes is demonstrated by the actuality of 65%, which falls into the *sufficient* category. However, the learning outcomes of the workshop participants have not met objectives of the workshop itself. The participants do not have complete understanding that practicing frequently is crucial to improve TOEFL scores. The causative factor for this discrepancy is the lack of exercising by participants either in doing exercise class or doing exercise independently. Lack of practice will lead to a lack of self-confidence and low motivation when answering test questions (Yogawati and Widhihastuti, 2019). This may happen because the strategy to answer questions that has been taught has not been operated optimally.

Contingency

The evaluation results indicate that the relationship between *antecedent* and *transaction*, *transaction* and *outcome*, and the relationship between *antecedent*, *transaction*, and *outcomes*, in terms of both intensity and observation, are all in the *sufficient* category. The findings are consistent with what was observed and supported by interview data indicating that some instructors still lack understanding of how to develop a suitable TOEFL lesson plan that adheres to the guidelines provided by the organizers. Besides, there are still instructors who have not carried out learning in accordance with the developed lesson plans. Some instructors are still lacking in exploring learning media, especially those related to latest technology that can support the success of TOEFL learning. This has an impact on less optimal learning outcomes carried out. The findings demonstrate that there is a relationship between the planning, implementation, and TOEFL learning outcomes in the TOEFL workshop program.

The adequacy of the TOEFL lesson plans made by the instructors falls under the sufficient category, indicating their ability to plan the learning process. However, this is

influenced by some instructors who lack an understanding of how to prepare such lesson plans, leading to difficulties in implementing the learning designs in class. As a result, the implementation of the learning process in the classroom is not optimal.

The learning implementation reflects the instructor's ability to manage the learning process, but it did not meet the standard process. As a result, the training participants were less engaged in learning, and it became difficult to execute the planned activities in class. The quality of the lesson plan prepared by the instructor affects the implementation of learning in the class, and if a good lesson plan is prepared, it will have a positive impact on the learning outcomes of the workshop participants.

The fact that TOEFL learning outcome is in the category of *sufficient* indicates that TOEFL learning plans and implementation is not fully in accordance with the guidelines set by the organizers, in this case is the UMB Language Center. Accordingly, it was found that there were still many workshop participants who had not received maximum TOEFL score.

CONCLUSION

The study draws the following conclusions: (1) The TOEFL learning plan falls in the *sufficient* category (69%), and the analysis reveals a discrepancy between the instructor-made lesson plan and the TOEFL lesson plan guideline. (2) The implementation of TOEFL learning falls in the *sufficient* category (60%), and the analysis reveals that it doesn't comply with the standard TOEFL workshop learning process. (3) The learning outcomes of the workshop are not all achieving the highest score, with an actual achievement of 65%, which is also in the *sufficient* category. (4) A contingency exists between planning, implementation, and TOEFL learning outcomes.

Based on the research conclusions, the followings are suggested: (1) TOEFL instructors should always improve their abilities in terms of learning strategies; (2) The implementation of learning should be well designed so that it can encourage the implementation of active, effective and enjoyable learning; 3) TOEFL instructors should try to use time allocated as effectively as possible for the smoothness of the TOEFL learning evaluation process.

The following recommendations are made based on the congruence and contingency analysis. First, good lesson planning needs to consider goal achievement through use of appropriate learning strategies suit workshop participants. Second, the implementation of learning should refer to the learning plan guidelines set by UMB Language Center. Third, TOEFL learning outcomes should be obtained by workshop participants through a learning process that increases time for practice.

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>
- 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>
- Iftanti, E & Maunah, B. (2021). Belief, Perception, and Challenges of Non English Department Students to Learn English in Globalization Era. *Indonesian EFL Journal*. 7(1), 1–12. <https://journal.uniku.ac.id/index.php/IEFLJ/article/download/3985/2452>
- 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.
- Esfandiari, M. R., Riasati, M. J., Vaezian, H., & Rahimi, F. (2018). A quantitative analysis of TOEFL iBT using an interpretive model of test validity. *Language Testing in Asia*, 8(1), 1-13.
- Esfandiari, M. R., Riasati, M. J., Vaezian, H., & Rahimi, F. (2018). A quantitative analysis of TOEFL iBT using an interpretive model of test validity. *Language Testing in Asia*, 8(1), 1-13
- Hamm, R.W. (1985). A Systematic evaluation of an environmental investigation course. (Doctoral dissertation. Georgia State University). ERIC Document. Reproduction Service No ED-256-622
- 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>
- Kaufman, R., & Thomas, S. (1980). *Evaluation without fear*. New York London: New Viewpoints
- Wang, Y. (2019). The impact on Instructors' Course Content and Teaching Methods. *The Electronic Journal for English as a Second Language*. 23(3), 1–18. <https://files.eric.ed.gov/fulltext/EJ1234792.pdf>
- Mardapi, D. (2017). *Pengukuran, penilaian, dan evaluasi pendidikan* (Revised ed). Parama Publishing.
- 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>

- Phillip, D. (2006). Longman Introductory Course for the TOEFL Test: The Paper Test. London: Pearson Education, Inc.
- Provus, M., Malcolm. (1969). The discrepancy evaluation models. An approach to local program improvement and development. Pitaburgh Public School.
- 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
- Stake, R E. (1967). Forward technology for the evaluation of educational programs. In R W Tyler, R M Gagne, & M Scriven. (Eds). *Perpectives of curriculum evaluation*. (pp.1-12). Chicago: Rand McNally.
- Stake, R E. (1967). The countenance of educational evaluation. *Teacher's Coolege Record*. Vol. 68, no:7.
- Stake, R E. (1967). The countenance of educational evaluation. *Teacher's Coolege Record*. Vol. 68, no:7.
- Stake, R E. (1977).The Countenance of educational evaluation.In A.A. Bellack & H.M Kliebard. Eds 1. *Curriculum and evaluation* (pp. 372-390).Berkeley,CA McCutehan.
- Stufflebeam, D. L. (1973). Educational evaluation: theory and practice. In B.R. Worthen & J.R. Sanders (Eds), *Evaluation as Enlighbtenment for Decision Marking* (pp.3-5). Boston: KluwerNijhoff Publishing.
- Stufflebeam, D. L., & Shinkfield, A.J. (1984). *Systematic evaluation a selfinstrucional guide to theory and practice*. Boston: Kluwer-Nijhoff Publishing.
- Wood, B.B. (2001). Stake's Countenance Model: Evaluating an Enveronmental Education Proffessional Development Course. *The Journal of Environmental Education*, Vol.32, No.2, pp. 18-27
- Yogawati, N. D., & Widiastuti, 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>

How to Cite (APA style):

Maryansyah, Y & Hamzah, S & Hadiwinatno. (2023, May). An Evaluation on TOEFL Workshop Program Using Stake's Countenance Model. *Exposure: Jurnal Pendidikan Bahasa Inggris*, 11(2), 32-46. <https://journal.unismuh.ac.id/index.php/exposure/article/view/97231>