

AN EVALUATION OF ENGLISH SPEAKING ACCELERATION PROGRAM AT IKIP PGRI PONTIANAK

Dedi Irwan

English Education Study Program
Language Faculty of STKIP-PGRI Pontianak
e-mail: dediirwan@hotmail.com

Abstract

This research aimed at finding out the discrepancy existing between the program standard criteria of English speaking acceleration program and its implementation. In order to run the evaluation, program evaluation was implemented at this research, while the model used was the Provus' discrepancy model. The research subjects involved were the stakeholders in charge at this program such as the key people in charge, program staffs, lecturers and program participants. The program implementation discrepancy was investigated by measuring the discrepancy at program input implementation, program process implementation and program output implementation. Program output variable was investigated from two aspects: interim output and terminal output. The data analyzing technique used was qualitative analysis enhanced by the quantitative measurement. The findings of this research are the program standard criteria and discrepancy score for the input implementation, process implementation and output implementation of the program. The implementation of input criteria is 81.33%, the process criteria is 4.77% and at the ultimate output is 85%. The terminal output evaluation finds the improvement of vocabulary mastery level up to 22.28%, the improvement of speech accuracy is 31.37% and the improvement of the pronunciation is 29.9%. Besides, it is also found that the decrease of the word rate from 2.095 at the pretest which falls to 1.939 at the posttest.

Keyword: Program evaluation, Provus' discrepancy approach, English speaking acceleration program.

Abstrak

Penelitian ini bertujuan untuk mengetahui adanya discrepancy antara kriteria standar program Akselerasi kemampuan berbicara Bahasa Inggris terhadap penerapannya di lapangan. Jenis penelitian ini adalah evaluasi program dengan menggunakan model kesenjangan Provus. Objek penelitian ini adalah setiap stakeholder yang terlibat dalam program seperti penanggungjawab program, staf, dosen dan peserta program. Penginvestigasi kesenjangan penerapan program diantaranya dilakukan dengan menginvestigasi Kesenjangan pada aspek input program, aspek proses program dan aspek output program. Teknik analisis data dilakukan dengan pendekatan kualitatif yang dikuatkan dengan analisis quantitative. Hasil penelitian ini adalah berupa kriteria standar program dan nilai kesenjangan yang terjadi pada penerapan kriteria standar input, proses dan output program. Penerapan pada input program adalah sebesar 81.33%, pada proses program adalah 4.77% serta pada output jangka panjang sebesar 85%. Sedangkan untuk output jangka pendek ditemukan peningkatan kemampuan penguasaan kosa kata sebanyak 22.28%, peningkatan pada akurasi berbicara sebesar 31.37%, peningkatan pada pengucapan sebesar 29.9%. Disamping itu ditemukan juga penurunan jumlah kata yang diucapkan dalam setiap detik dari 2,095 kata pada pretest menjadi 1,939 kata pada posttest.

Kata Kunci: Evaluasi Program, Pendekatan kesenjangan Provus, Akselerasi kemampuan berbicara bahasa Inggris.

EVALUATION BACKGROUND

The massive improvement of technology nowadays brings the human being a borderless world where every single place in this world get connected easily one to another. This condition sets the people to be in a high competitive life. To be able to survive in such condition, a global cooperation with a great deal of international communication is really in need. English, as an international mean of communication, plays a key role. By mastering English would mean being able to survive in such competitive condition. On the other words, as global citizen like now, competitive people cannot be separated from the English mastery.

To be able to help student master English faster, the Study Program officer had been attempting numbers of strategies. One of the relatively new strategies was by creating elite technocrats in speaking through the English Speaking Acceleration Program as differentiated by Ganiron (2013, p. 28) from the traditional way. Learning English Speaking Acceleration Program, which was designed as a pilot project at STKIP PGRI Pontianak, was implemented to a small group of selected students to have a special teaching learning system. This program was administered by English Education Study Program cooperated with Sinka English Training Centre (SETC) Singkawang, West Kalimantan. English Speaking acceleration program is a program designed to facilitate the outstanding students of English Education Study Program of STKIP PGRI Pontianak, especially in English speaking skill, to have a special learning system.

In each year, there are maximum 20 students who are selected through a highly –competitive selection process. Those selected students are then sent to join this Program in Singkawang for one semester. In the learning process, the Participants are taught during day and night in order to develop their ability in speaking. Compared to the regular students, the selected Participants would spend more time to learn everyday during the program. By which it is expected that the Participants would be able to speak English more fluently than the regular Participants could do.

This evaluation is a quantitative descriptive enhanced by qualitative data. The quantitative descriptive was used to measure the discrepancy happened in

each stage, While qualitative descriptive was used to describe specifically any factors constrain the implementation of program by investigating the program's internal and external consistency. Based on the main problem of this evaluation, the purpose of this evaluation in general was to evaluate how far the speaking acceleration program has been implemented based on its objectives that were formulated by the program stakeholder. In detail, the purpose of this evaluation was to define the program's criteria standard, evaluate the program implementation and evaluate the program's output. The purpose of this evaluation was under the belief that "*an educational intervention, academic acceleration is decidedly effective for high-ability students*". Colangelo, (2010, p.2). In a more qualitative comparative study, Conrad quoted from Wlodkowski (2003, p.2) found that intensive courses became rewarding and powerful learning experiences when certain attributes were present

EVALUATION METHOD

This evaluation is a quantitative descriptive enhanced by qualitative data. The quantitative descriptive was used to measure the discrepancy happened in each stage, While qualitative descriptive was used to describe specifically any factors constrain the implementation of program by investigating the program's internal and external consistency.

The type of evaluation used was program evaluation that is "*a process of focusing attention on the process of education using professional judgment and developed standard for education programs.*" Olaitan (1996, p.1). Moreover, Drucker (1977, p.1) emphasizes that evaluation is a watchdog of program management. It ensures that standard can be used for assessing program performance and students productivity.

The model of discrepancy used was Provus' discrepancy model. Provus' discrepancy model is "*Methods which are traditionally used to evaluate products, such as the simulator-based evaluation, do not provide a systematic and comprehensive means for identifying flaws that may be contained within the product.*Regan et.al (2001, p.6).Provus (1969, p. 9) explains that this type of

evaluation at its simplest level may be seen as the comparison of performance against a standard. McKenna (1981, p.14) explains that Provus' model offers five steps of doing discrepancy evaluation.

The steps start from establishing program design standard, and then is continued by planning evaluation using the discrepancy model, collecting information on performance, identifying discrepancies, and then alter performance and/or alter standard. *"This is argued as a barrier to learning patients' behavioral patterns and understanding program performance."* Mills (2010, p.516)

This research was conducted at two places that were at English Education Study Program of STKIP PGRI Pontianak as the administrator and at SETC Singkawang as the managing institution of this program. This Evaluation was conducted from December to June 2014. A set of preparation was done during December to early of February. The data were collected from mid February to the early of May. The data analysis was conducted during May. The subjects of this evaluation were stakeholders involved in the English Speaking Acceleration Program of English Education Study program including study program officers, program staff, lecturer, and program participants. There are 16 participants from the first semester of 2013 class, and 14 from the second semester.

The data collected in this evaluation were in form of qualitative and quantitative data. The quantitative data were used to measure each discrepancy, while the qualitative data used to describe the factors caused the discrepancy. In collecting the data, the evaluator passed through some steps. In designing the program criteria standard, the evaluator involved program administrators and every key person who was in charge in this program. This process aimed at developing the English Speaking Acceleration program design criteria. The data were collected by conducting interview and doing document study. The data that were collected through this stage was the data about the program criteria standard for program design including program input, program process and program output. The interview was done toward the key people in charge of this program

To investigate the implementation of the program, the evaluator involved two aspects that was internal and external consistency. Internal consistency deals

with the readiness of the program components. The readiness here refers to the key aspect of this program that hold the key role in the success of this program's objective. Some of the key aspect are the participants' characteristics, Participants' activities, staff qualifications, staff' activities, supporting administrative, and supporting media. The external consistency refers to the program's compatibility. It involved a study of the compatibility of the program operative in the entire campus system at STKIP PGRI Pontianak in general. This aspect investigates the effect of the existence of this program toward the entire program available at the campus level

In investigating the internal and external consistency, the evaluator used inter rater judgment that consisted of participants, program administrator and observation conducted by the evaluator. The raters stability was measured by using Intra-class Correlation Coefficients (ICC), the score got was 0.63 or 63% that was considered as **enough**. In measuring the discrepancy between the program criteria standard against the implementation, out of the limitation the evaluator had, the evaluator only involved two source of information in order to measure the discrepancy happened namely through the participants and through observation.

EVALUATION FINDINGS AND DISCUSSION

Based on the evaluation purpose, this evaluation aimed at developing the program criteria standard, investigating the program internal consistency, investigating the program external consistency, and investigating the discrepancy of program criteria standard against its implementation.

Program input was seen from five aspects such as program participants, staff involved, lecturer involved, administrative support, and media available. Program participants covered participants' selection criteria and focus skill to develop. The selection criteria for the participants were conducted in two parts namely speaking fluency and vocabulary mastery. Interviewing Participants tested the speaking skill, while the vocabulary mastery was tested by using a written test that capable of showing the Participants' vocabulary mastery. All of the tests were conducted systematically. The documents used were well documented. The

selection was conducted in competitive way.

Each variable above was measured by a set of indicators. The following Figures 1 describe the score of discrepancy for each indicator.

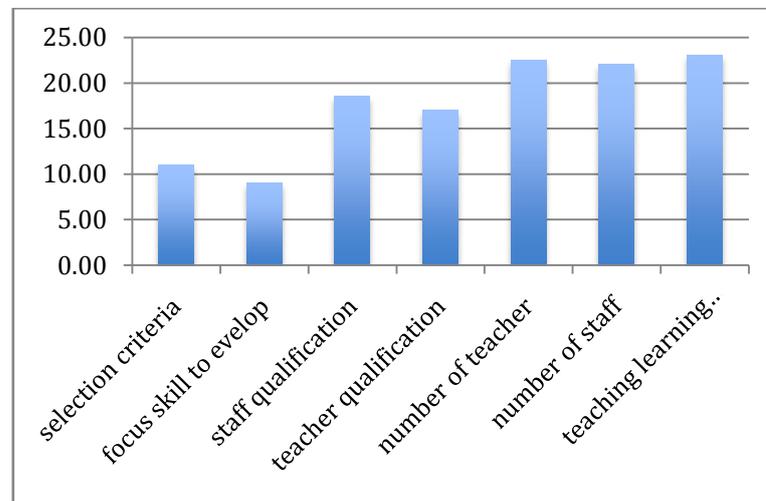


Figure 1. Discrepancy score for each input indicator

From all those indicators, number of lecturer, number of staff and teaching learning media placed at the big three reaching 23%, while the lowest discrepancy was at selection criteria and focus skill to develop. Based on the collected data about focus skill to develop, this program focused on developing the speaking skill. The teaching materials were developed with the basic skill of speaking. Most activities of the participants were aimed at practicing the Participants' speaking ability.

Staff involved in this program were selected from the campus' staff. The staff including the study program chairperson, study program secretary, and study program staff. Besides, there are also some other staff hired from the other part, most of them work as the visiting lecturer and kitchen lady. While the lecturers were selected from the study program. There were eight lecturers involved. The educational background of the lecturers were master degree, even there was one lecturer having doctoral degree.

The program process covers participants' regular activity, staff function

and the lecturer role in helping the participants. In daily activity, the participants, during the program, works on two curriculum. In the morning the Participants learn by sung the campus curriculum just like the non-participant Participants while during the afternoon till the night the learn by using the acceleration syllabus. Within the acceleration syllabus the participants learn in various technique of learning namely speech delivery, storytelling, singing a song, playing drama and teaching practice so called 'buddy'.

To support the process, the staff should be responsible to the preparation of supporting media and preparing the administrative services while the lecturers are responsible at teaching by using common way in the morning classes and lecturers become a facilitator for the participants in a contextual teaching learning class.

The output of this program was divided into two types; ultimate output and the ultimate output. The ultimate output describes the direct short time effect of the program toward the participants while the ultimate output is the long run effect gained by the participants. The ultimate output of this program was helping the participants to increase their speaking fluency and enriching their vocabulary mastery. While the ultimate output of this program is that to produce the peer tutor, produce the ambassador of English Education Study Program and preparing their future career.

The activity of the participants during the program was divided into two parts: learning by using the regular syllabus during the morning and the acceleration syllabus from afternoon to night. During the acceleration syllabus, the participants learn to develop their vocabulary mastery and speaking fluency by using different techniques such as speech delivery, drama playing, singing a song, teaching practice (Buddy), and telling story. Below is the discrepancy score for each indicator

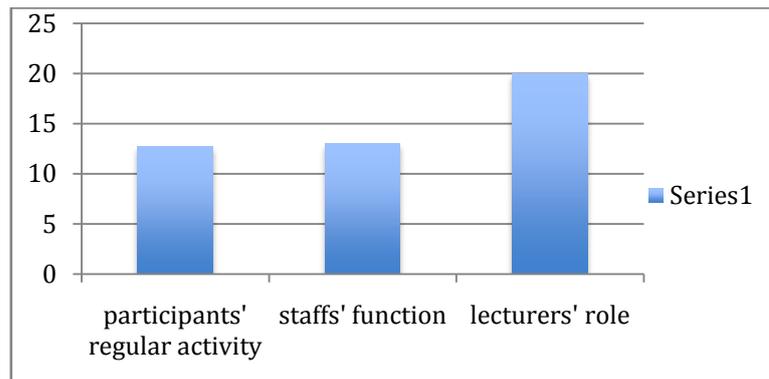


Figure 2: Discrepancy score for process indicators

Based on the data above, the program staff, program lecturers and the participants believe that all those activity had been relevant toward the program goal. However, the evaluator agreed just a half since at the implementation the evaluator found some problem in each activity. The problem existed especially during the morning time where they learn the common syllabus. The problem mostly about the number of lecturer absence was still high. Beside, for the acceleration syllabus, the evaluator did not find any specific lesson plan, or detail procedure for each activity. This fact, of course, affects the goal achievement of the program for there is no tangible and systematic activity guidance for the participants during the program.

The lecturer involve in this program were various, starting from diploma to doctoral degree. The staff fully believes that the qualification had been enough. The other reason that convinced them was that all the lecturers had linear education background toward the English speaking teaching. The Participants, in this case, agreed the staff' judgment with some complains. The complaint was about the lack attendance intensity of the S2 lecturer at the program. They found that in certain time, the lecturer sent to join the program was the s1 ones so that, according to them, it was disappointing. However, according to the evaluator observation, it was true that on the list the qualification of the lecturer had been enough. But unfortunately, during the process of the acceleration curriculum, most of the lecturers were not involved. The lecturers from English Education Study Program only involved during the morning class.

The program staff, the participants and the evaluator, in this case, agreed that there was a problem at the staff duty description. Though it was not major problem but it could interfere the program goal achievement. The problem existed at the number of the lecturer involved during the acceleration program. Most of the staff should share their time with the other program of the campus so that the regular schedule often ignored. This situation caused the bias to the description of the duties.

In case of the administrative support, participants of this program fully agree at this program had been enough. The reasons they made as the foundation was that during the program they did not find any problem related to the administrative things. This judgment was agreed by the staff, but some note came up from the staff. The note was about the number and intensity of the staff available at the program. The place where the program administered caused the problem for the staff to join regularly. The evaluator was at the same boat in this case. The hectic schedule and the fact that they should do three hours riding to reach the place of the program administered cause them incapable of joining the program regularly. However, the evaluator and the staff believe that it was not a major problem since the duties could still be covered by some staff staying at the program though it was not optimum.

Pretest-posttest of vocabulary mastery level involved 14 Participants. The posttest was consisted of 93 items and was conducted right after the program finished. The program took four months. The data collected for this step is presented below.

Table Score Description of Vocabulary mastery level pretest and posttest

Component	Pretest	Posttest
Mean	50.92 (54.78%)	71.71 (77.07%)
SD	18.16 (19.49%)	12.64 (13.50%)
Min. Score	24 (26%)	35 (38%)
Max. Score	76 (82%)	86 (92%)

Based on the data collected, the minimum score for the pretest, as shown at the figure above, was 24 and the maximum score was 76. While for the posttest, the minimum score was 35 and the maximum score was 86. From the statistic, it explicitly appears that this program improved the participants' vocabulary mastery level. The lowest score increased up to eleven points while for the highest score increased up to 10 points. This data describe that both lower achievement and high achievement participants could develop themselves after joining this program. This is supported by the mean score that increases as much as 20.79%. Furthermore, the knowledge of the participants also becomes more homogeneous as implied by the Standard Deviation that decreases from 18.16 to 12.64.

The next indicators noted from the data to support the participants' improvement in vocabulary mastery, that is also the most interesting finding, was about the gain score. As shown at Figure above, the lowest gain score was 6 (6%) and the highest gain score was 50 (54%). The mean score of the gain score was 20.78 or 22.28%. When the score was compared to the score needed to reach the maximum score (100%) that is 45.22%, the implementation of this program was as much as 49.27% which meant that the discrepancy happened as much as 50.72%. Based on the number of discrepancy identified, the implementation of this program was categorized as **Poor**. Based on the score distribution presented at the Figure, the evaluator split the score to four classes of score range in order to ease the analysis as shown the following Figure. Based on the score distribution above, the following score range is presented.

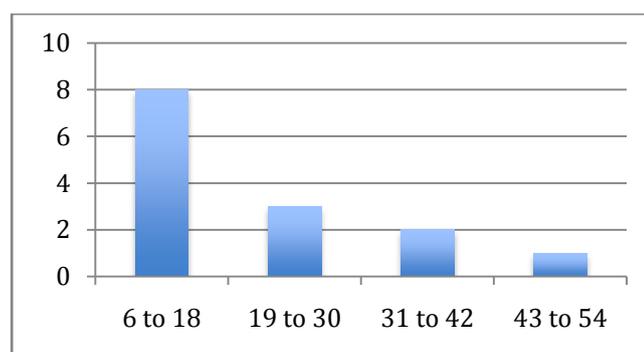


Figure 3. Gain Score Percentage Range

From the Figure, it is found out that the highest number of score was at the range of 6 to 18 and it keeps decreasing to the next three ranges consistently. This distribution of data grew the curiosity of the researcher to find out more about the score distribution. Therefore To investigate the gain score distribution, the evaluator looks deeper toward the relation between the pretest score and the gain score. The comparison is presented as below.

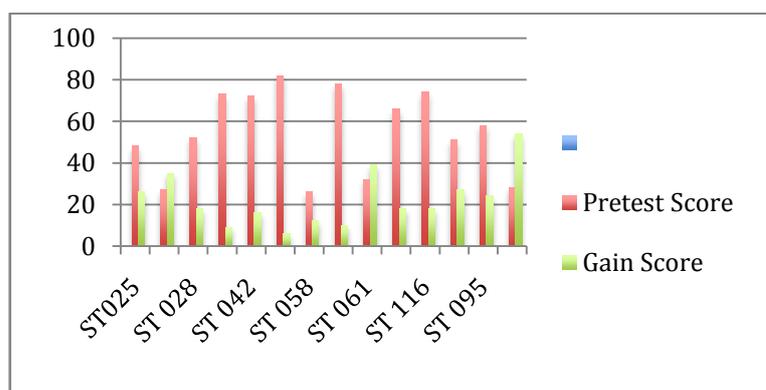


Figure 4. The comparison of pretest score against the gain score

The three highest gain score were 54, 39 and 35 owned by the ones whose pretest was 28, 32 and 27 in respective order. While the three least gain score was 6, 9 and 10 belonging to the one with the pretest score of 82, 73 and 78 respectively. It shows that the participants with lower scores could develop themselves much more than the ones with the high score of pretest. The statistic above was then simplified by dividing the spread of the gain score based on two categories of pretest score, that is more than 50 and less than 50. The data description is shown as below.

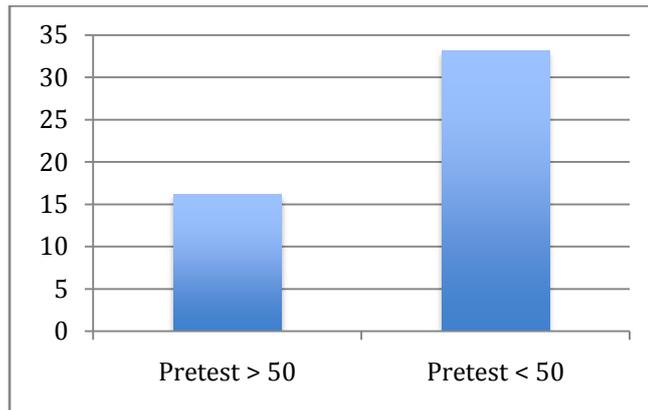


Figure 5. The Gain Score Mean Based on Pretest Score

Based on the Figure it is found out that the participants whose pretest were score less than 50 could increase their vocabulary mastery more than 50% than the one whose pretest score more than 50 could do. In average, the participants whose pretest score less than 50 could develop their vocabulary mastery up to 33.2% while the ones whose pretest score more than 50 could only develop their vocabulary mastery up to 16.22%. Based on leveling criteria, the improvement of the ones whose pretest was more than 50% at the level of very poor since the discrepancy happened was at the range of 75-100% while the improvement of the one whose pretest less than 50% was at the level of **poor** since the discrepancy identified was at the range of 26-50%.

The last indicator counted in measuring the Speaking Skill is the Words rate. The word rate was counted by calculating the utterances produced in each second. Based on the data collected, below presented the word rate average in pretest and posttest speech.

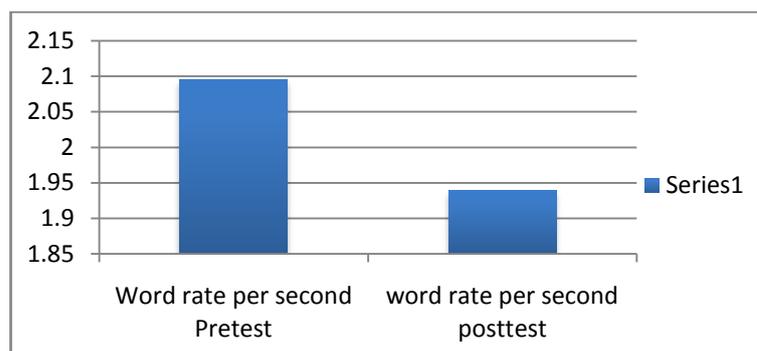


Figure 6. Pretest-posttest word rate average

Based on the data collected as shown through the Figure above, there is the decreasing of word rate between the pretest and the posttest. The rate decreased from 2.09 at the pretest becomes 1.93 words per second.

CONCLUSION

Based on the data collected, there is nine conclusions of this evaluation are concluded. First, input standard criteria of this program are (1) Students with minimum GPA of 3.5 who passed the vocabulary mastery level test and oral test, (2) Lecturer and staff with the qualification of doctorate and master, (3) administrative support involving eight staffs and lecturers, (4) natural context teaching-learning media.

Second, program process standard criteria are (1) the participants learn using the regular syllabus during the morning and using the acceleration-learning syllabus for the rest of the day, (2) the staff is responsible to prepare the supporting media and administrative services, (3) the lecturers were to teach and facilitate the participants.

Third, output of this program was divided into two (1) interim output; improving the participants' English speaking fluency, skill of English Speech and Mastery of English vocabulary better than the non-participants would do and (2) ultimate output; Producing ambassadors for English Education Study Program and producing peer tutors for the non-participant students.

Fourth, score for the implementation of input standard criteria according to

the program participants was 499 out of 672 that equal to 74.25%. It means that according to the participants only 74.25% of the input standard was implemented and left 25.75% discrepancy. Among all indicators on this variable, the attendance of lecturer, number of staff and teaching-learning media were the lowest among all so that they need to be improved.

Fifth, the score for the implementation of the process standard criteria according to the participants was 245.33 out of 288 that equals to 88.30%. it means that according to the participants there were 88.30% of the process standard criteria were implemented, it left 11.70% discrepancy. From all indicators involved, lecturers' role in helping the participants pursuing their goals in this program still needed to be improved.

Sixth, Interim output criteria implementation was divided into two; participants' vocabulary mastery level and the speaking skill. In mastering vocabulary, the participants could improve their mastery from averagely 50.92 in the pretest to 71.71 in the posttest that equal to 49.40%. It means that based on the test result, it is need 50.59% more improvement for the participants to reach maximum score in vocabulary mastery test. Unfortunately, the improvement was gained the most only by the one whose pretest below 50% that improved from averagely 35 in pretest to 63,14 in the posttest, while for them whose pretest score more than 50% could develop their mastery in vocabulary only for 16.22% or improved averagely from 66.85 in the pretest to 80.28 in the posttest. In speaking skill, the participants get increased in their speaking skill after joining this program in different amount for each sub skill. The speech accuracy increased for 31.37% that means that there still needed 68.63% of improvement for the participants to reach the maximum score. For the pronunciation aspect the participants got increased only for 29.9%. It means that based on the result of this test the participants are in need 70.10% of improvement in order to reach the maximum score.

Seventh, average score for the ultimate output according to the participants was 235 out of 288 that equals to 81.59%.It means that according to the participants, the implementation of the ultimate output standard criteria was

81.59%. In this case, the participants were still unsure a hundred percent that they could be peer tutor and study program ambassador.

REFERENCES

Colangelo, N et.al. (2010). *Developing an Academic Acceleration Policy*. Journal of Advance Academic Volume 21, Number 2 Winter 2010 pp. 180–203

Drucker, D. F. (1977). *Management: management park, responsibilities and practices*. New York: Harper's College

Ganiron Jr, T.U. (2013). *Application of Accelerated Learning in Teaching Environmental Control System in Qassim University*. International Journal of Education and Learning Vol.2, No.2, pp.27-38

McKenna, C. (1981). *Making Evaluation Manageable*. Journal of Extension september/october.

Mills, J.A. et.al.(2010). *Encyclopedia of case study research*. Volume 1. Washington DC: Sage Publication.

Provus, M.M. (1986) *The Discrepancy evaluation model. An approach to local improvement and development*. USA: U.S Department of health, education and welfare

Reagan, et.al.(2001). *Provus' discrepancy evaluation of the drivesmart notice driver CD- Room training product*. Monash University Accident research Centre Journal.

Wlodkowski, RJ. (2003). *Accelerated Learning in Colleges and Universities*. New Directions for Adult and Continuing Education Journal, No. 97, 5-15.

