

Evaluation of Pre-Departure English Training Program

Ali Saukah

Abstract: The program evaluation reported in this article covers three batches of participants, from 1996 through 1997, sponsored by the Directorate General of Higher Education projects to prepare faculty members of teacher training institutions for overseas studies. The result of the program evaluation indicates that the highest possible learning outcomes could be achieved when the initial English competence required for participating in the program was at least at the Pre-Advanced level. The criterion validity of the TOEFL-Equivalent test developed by the program was assured, and the test could, therefore, be used as a good predictor of the International TOEFL. Other recommendations are also given for the improvement of the planning and implementation of the program in the future.

Key words: program evaluation, English training, TOEFL, overseas studies

Pre-Departure English Training Course (PDETC) program has been conducted for more than fifteen years in the Graduate Program of IKIP MALANG. As the name indicates, the program has been primarily designed to prepare the participants for overseas advanced studies. The main objective of the program is to help the participants reach a level of English language performance as indicated by their TOEFL scores. To achieve the objective, a curriculum has been developed which is essentially focused

Ali Saukah is a Senior Lecturer at the Department of English Education, IKIP MALANG.

on the mastery of all the four language skills. A study on the relationship between English language proficiency and academic performance has revealed that sufficient English language proficiency to undertake the academic program at the universities using English as the medium of instruction should be given foremost consideration (Johnson, 1988).

Evaluation, in this context, should be defined as a systematic attempt to examine what happens in, and as a result of, the Pre-Departure English Training Program in order to find the basis for judgments and decisions about this program. In conducting the evaluation, a number of approaches have been suggested in the program evaluation literature. Brown (1995), for example, introduces four different approaches to language program evaluation: product-oriented, static-characteristic, process-oriented, and decision-facilitation approaches. Each of the categories, however, seems to focus only on specific issues of the program. A more comprehensive framework of program evaluation which takes care of most, if not all, of the important issues in the whole program should be developed. Such an approach should address the full range of concerns of the program (Lynch, 1990).

On the eyes of the clients, the primary concern is whether they can decide to continue collaborating with the Graduate Program of IKIP MALANG in preparing their human resources for overseas studies. The primary concern of the program developer, however, is whether things need to be revised or improved. This is in line with the idea of Stufflebeam & Shinkfield (1985) that the most important purpose of evaluation is not to prove but to improve.

The evaluation reported in this article covers three batches of PDETC funded by the Directorate General of Higher Education projects. The first batch was held from September 11, 1995 to February 10, 1996, with funds from Academic Staff Training (AST) Project as part of the preparation for Secondary School Teacher Development (SSTD) Project. The second batch was carried out from July 1, 1996 to December 20, 1996, with funds from the AST project which supports the division of the SSTD Project related to sending academics of teacher training institutions for further study overseas. The third training program was carried out from June 23, 1997 to December 12, 1997, with funds directly from the SSTD Project. The three English training programs were held by the Postgraduate Program of IKIP MALANG. The main objective of the PDETC is to

improve the participants' proficiency in English as an effort to fulfill the formal requirements for enrollments to universities abroad which are approved by the Directorate General of Higher Education, the Department of Education and Culture.

The evaluation of the intensive English Training Program is aimed at providing input for the Secondary School Development Project, as the client of the program evaluation, in order to determine: whether the intensive Pre-Departure English Training Course needs to be continued in the coming years; the extent to which the test results given by the teaching staff of the ETP correlates with the international TOEFL; if it is considered that the program needs to be continued, what improvements or revisions are required to increase the efficiency, efficacy and the quality of the planning, programming and implementation of the Intensive English Training Program.

METHODS

There were 50 participants in the first batch, selected from 340 candidates for the selection carried out at 17 Teacher Training Institutions. The participants for this first batch of training were those who had a minimum TOEFL score of 420, and were not over 35 years of age.

The participants of the second batch were initially 75 people selected from 330 applicants who took a selection test at 24 Teacher Training Institutions. The minimum TOEFL score of the participants of this second batch was 410, and the maximum age was also 35. Four participants then withdrew from the program, this making the final number of participants 71.

The number of participants for the third batch was 43. They were selected through the Pra-S2 (pre-master's degree program) selection conducted at 29 Teacher Training Institutions. These participants were selected from the list of participants for the Pra-S2 program selection who were not admitted into the program and were still within the age limit. Using that criteria as the basis of consideration, there were finally 43 participants selected with a minimum TOEFL score of 330.

Most of the data needed for the program evaluation were collected by way of documentation, i.e. by collecting documents related to (1) the

objectives of the intensive PDETC and the background of those objectives, (2) the characteristics and qualifications of the participants and how they were selected, (3) the realization of the intensive PDETC as reported by the committee, and (4) the English test scores achieved by the participants during the training program. These scores were obtained from both the TOEFL-equivalent and the international TOEFL.

The data obtained for this evaluation were analyzed both quantitatively and qualitatively. Data in the form of scores-either from the result of the English test or from the questionnaire on the implementation of the training program were analyzed using statistics. The results of the analysis were mostly focused on the predictive validity of the tests developed by the training program committee towards the results of the international TOEFL and towards the efficacy of the training program in terms of the participants' proficiency in English. Data in the form of verbal explanations and answers were analyzed qualitatively.

Based on the results of the analysis, recommendations would be given to the Secondary School Teacher Development Project to enable them to decide on (1) whether the intensive English training Program should be continued in the future and (2) the improvement and revisions that need to be done in order to increase the efficacy, efficiency, planning and implementation of the intensive Pre-Departure English Training Course if it is continued.

RESULTS AND DISCUSSIONS

The results of the evaluation and the discussions will be presented below covering: (1) the participants' initial and final proficiency (at the start and end of training), (2) the validity of the test, (3) the curriculum of the program, (4) the teaching-learning process, (5) instructors, (6) facilities, and (7) grouping of participants.

Initial and Final Proficiency of Participants

The participants' proficiency at the beginning and end of the training program is shown in Table 1. The table shows that the participants' average English proficiency at the start of the program for all groups was at the level of Pre-Advanced with scores ranging from 373 to 590, and the

highest average score was possessed by the English Group (EG). However, there was a significant difference in the maximum scores among the four groups. The highest initial score (590) was possessed by the English Group, followed by Hard Sciences Group (577) and the Social Sciences Group (517). The Other Language Group had the lowest initial score (493). Besides having the highest initial score, the English Language Group also had the biggest number of highest proficiency (Special Advanced). Out of the five Special Advanced scores, four belonged to the English Language Group, which is undoubtedly in accordance with the expectations.

Table 1 The Participants' Initial English Competence of all Batches as Indicated by their TOEFL-Equivalent Scores

Fields of Study	Total		Mean	Range	SP-ADV		ADV		PRE-AD		INT	
	n	%			n	%	n	%	n	%	n	%
English	56	34	482	407-590	4	7	15	27	32	57	5	9
Language	14	8	438	407-493	—	—	—	—	7	50	7	50
Social Sc.	46	27	442	373-517	—	—	4	9	28	61	14	30
Hard Sc.	50	31	455	413-577	1	2	5	10	36	72	8	16
Total	166	100	459	373-590	5	3	24	15	103	62	34	20

Notes: The fields of study include: (1) English Education and Literature (English), (2) Language Education other than English (Language), (3) Social Studies Education and Education in general (Social Sc.), and (4) Math, Physical Sciences, and Technology Education (Hard Sc.)

At the end of the training program the highest average score was achieved by the English Language Group (516), followed also by the Hard Sciences Group (481), and the Social Science Group (447). The Other Language Group also came last (441). The order was exactly the same as that at the start of the program. It indicates that initial proficiency plays an important role in determining the final score at the end of the training.

Table 2 The Participants' English Competence of all Batches at the End of the Program as Indicated by their International TOEFL Scores

Fields of Study	Total		Mean	Range	SP-ADV		ADV		PRE-AD		INT	
	n	%			n	%	n	%	n	%	n	%
English	56	34	516	433-603	9	16	27	48	20	36	—	—
Language	14	8	441	300-537	—	—	2	14	7	50	5	36
Social Sc.	46	27	447	350-530	—	—	5	11	26	56	15	33
Hard Sc.	50	31	481	387-597	1	2	13	26	31	62	5	10
Total	166	100	480	300-603	10	6	47	28	84	51	25	15

Besides showing comparison of the maximum average scores, the figures in Table 2 also demonstrate the highest level achieved by participants in each of Fields of Study groups. In this case, the highest score (603) was also obtained by the English Language Group, followed once again by the Hard Sciences Group (597). The subsequent order showed a difference, with the Other Language Group in the third place (537) and the Social Sciences the last (530).

The comparison between the average scores at the beginning and end of program can also show the participants' gain as shown in Table 3. The gain is reflected in the difference between the average initial score and the average final score for each group. The computation of the gain again reflected the rank which was basically consistent among the groups. The English Language Group came first (with a gain of 34 score points), The Hard Sciences Group second (with a gain of 26), the Social Sciences Group third (with a gain of 25), and the Other Language Group last (with a gain of 10).

Table 3 The Participants' English Competence of all Batches At the Beginning and End of the Program

Fields of Study	TOEFL Scores	N	Mean	Max.	SP-ADV		ADV		PRE-AD		INT	
					n	%	n	%	n	%	n	%
English	Initial	56	482	590	4	7	15	27	32	57	5	9
	Final	56	516	603	9	16	27	48	20	36	—	—
	Gain		34	13	5		12		-12		-5	
Language	Initial	14	438	493	—	—	—	—	7	50	7	50
	Final	14	441	537	—	—	2	14	7	50	5	36
	Gain		3	44	—	—	2	—			-2	
Social Sciences	Initial	46	442	517	—	—	4	9	28	61	14	30
	Final	46	467	530	—	—	5	11	26	56	15	33
	Gain		25	13	—	—	1		-2		1	
Hard Sciences	Initial	50	455	577	1	2	5	10	36	72	8	16
	Final	50	481	597	1	2	13	26	31	62	5	10
	Gain		26	20	0		8		-5		-3	

The increased proficiency of the participants furthermore can be traced from the greater number of participants who could reach a higher level of proficiency. As the TOEFL test indicates, there are six levels of English proficiency, i.e. - Elementary (with a score lower than 200), Pre-intermediate (201-350), Intermediate (351-425), Pre-Advanced (426-500), Advanced (501-550) and Special Advanced (551 or more). With the Advanced level as a basis, the increased proficiency of a group was indicated by the increase in the number of participants who could reach that level, and at the same time by the decrease of the number of participants in the lower levels. From these descriptions the same picture could be obtained about the four groups, with a consistent order: the English Language Group (with additional 5 Special Advanced and 12 Advanced), the Hard Sciences Group (additional 8 Advanced), the Other Language Group (additional 2 Advanced), and the Social Sciences Group (additional 1 Advanced). All the addition to the higher level was automatically followed by a reduction in the levels lower than Advanced-as shown in the table.

The only "abnormality" was the increase in the number of intermediate level in the Social Sciences Group (from 14 to 15 Intermediate).

The Validity of the Test

The test used to measure the initial proficiency of the participants was a TOEFL-Equivalent test because it was much less expensive than the International One. Besides, it was impractical to administer the International TOEFL for the program candidates as a selection instrument-by which the initial proficiency of most participants was measured-because they came from almost all teacher training institutions (LPTK's) throughout Indonesia. Their final proficiency, however, was measured by the International TOEFL administered at the location of the PDETC at the Graduate Program of IKIP MALANG, which happened to be one of the International TOEFL testing centers. The validity evidence for the use of the TOEFL-Equivalent test was provided by correlating the scores of the International TOEFL with those of the Equivalent One which were both administered at the end of each program.

Sufficiently high validity coefficients of the instrument can be inferred from the correlation coefficients as shown in Table 4.

Table 4 The Validity Coefficients of the TOEFL-Equivalent Test

	International-TOEFL ₁	International- TOEFL ₂
TOEFL-Equivalent ₁	0,5723	0,5186
TOEFL-Equivalent ₂	0,7315	0,7064
TOEFL-Equivalent ₃	0,6244	0,6382
TOEFL-Equivalent ₄	0,7286	0,7100

Note: All significant at $p \leq 0.000$, with $n=166$; each participant took

TOEFL-Equivalent test four times and International TOEFL twice.

The Curriculum

The evaluation of the curriculum is based on the notes, ratings, suggestions and inputs from the participants given anonymously at the end of the training program to a series of questions in the form of 'Program Evaluation'. The questions in that form were categorized into two groups. The first was closed-ended, where the participants had to give scores from 10 to 100 to every question. The evaluation was fully up to the participants. The second group consists of open-ended questions which the participants had to fill in with short answers or explanations. The value of a question was obtained by computing the average scores of all participants which were then converted into five categories: A (very good/very appropriate/very much needed/very satisfactory); B (good/appropriate/necessary/satisfactory); C (fair); D (poor); and E (very poor).

The answers to the open-ended questions were recorded as they were.

The question that can be categorized into a request for input and evaluation of the curriculum was 'In general, do you feel this training program has helped to improve your English proficiency?' The average score of the participants of the first batch (63 answers) was 80.19, equivalent to B (good/appropriate/necessary). The same answer was given by participants of the third batch (40 answers), which was 82.23, also equivalent to B. The participants of the second batch (69) gave an average score of 73.32, equivalent to C (fair).

Another question in the category of curriculum evaluation is 'In general, are the subjects given in the training program appropriate and useful?' The answer to that question was also B from batch I (average score 80.40), and batch III (average score 85.23), and another C from batch II (average score 74.58).

The answers to the open-ended questions on curriculum provided input and evaluation on various aspects related to the curriculum of the program. Some of the answers referred to curriculum in general, while others referred to some specific aspects. Most of the answers were positive, offering praises and satisfaction, while there were also some suggestions for improvement. These tendencies were present in all three batches.

Below are some positive expressions given as answers to the question 'What good points, if any, did you find in the training program? Give brief answer'. These answers can be categorized into curriculum evaluation, given in different categories with their direct quotes:

- (a) Curriculum in general:
- I got a lot of improvement from listening, reading, vocabulary, and grammar.
 - I feel I got a lot of improvement although this was the first time I joined an English course.
 - I got a lot of knowledge of English in detail, skills in using English and increased discipline.
 - (This program) has increased my ability in English, reinforced what I have already gotten, and provided me with new knowledge I have never learned.
 - I am more confident in using/practicing my English.
- (b) Some components of the curriculum:
- (I) can speak English better than before.
 - (I was) trained to catch the meaning of words spoken by the instructor quicker and I become accustomed to listening and speaking in English.
 - (I gained) knowledge and skills in writing and confidence in speaking in English.
 - The materials for writing course really helped in writing different types of essays.
 - The biggest advantage was from the subjects of grammar and listening.
- (c) The Wide Coverage of the curriculum
- Very useful strategy in tackling the TOEFL test.
 - ASO (Academic and Social Orientation) provided a lot of knowledge and experience about America.
 - Clear and extensive explanation and information about studying abroad.

The opinions given anonymously by the participants concerning the implementation of the program, including the curriculum, could include negative and critical views. The question was 'What weaknesses, if any, did you feel in this program? Give a very brief answer'. Most of the feedback reflected the lack of understanding in the basic strategies to improve their English proficiency, which cannot be achieved merely by practicing doing tests. Some of the feedback was as follows:

- The materials were not focused on the TOEFL so that the participants' scores could not increase easily.
- (The materials should be) purely TOEFL.
- The vocabulary lessons were not sufficient as in TOEFL.
- The TOEFL materials need to be added.
- There needs to be a standard syllabus.

Teaching Learning Process

Participants' positive responses toward the teaching and learning process, which constitutes the main activities of the training, were also revealed in their answers to the question concerning the strengths they had found. The responses cover the following points:

- Good management and selection of training facilities.
- Providing the participants with teaching models they would be able adopt later in their own institution.
- Good management, no absent lecturer without being substituted (only once, ASO class)
- Varieties of teaching techniques and methods were given, especially for teaching the four language skills.
- The lecturer explained in the target language and the participants got opportunities to speak in the language.

There were also some notes and comments showing the weaknesses, which were treated as feedback for better management in the future, for examples:

- Participants with lower competence needed more attention, especially for Listening.
- Independent learning conducted in the campus should have been provided.
- Some lecturers were too dominant, less interactive, and less open to participants' questions.
- There should have been more involvement of native speakers in Speaking.
- The program was too overloaded, no time to study at home.

Instructors

Participants' responses to the instructors as shown in their answers to the questions concerning the strong points of the instructors were mostly positive. These positive responses were mainly concerned with their teaching skills in relation with their academic qualification and teaching experience, in addition to their discipline and their interaction with participants. Some examples of their comments are:

- The instructors had great competence.
- Most of the instructors were friendly and patient
- The instructors had good competence, good discipline, and good teaching methods.
- The instructors seriously cared about the participants.
- The instructors were very keen, hardly absent/did not send substitutes.

In response to the question 'What are your comments and/or suggestions and general evaluation toward the instructors' teaching performance?' the participants evaluated the instructors quantitatively. Based on the same score used for other questions and on the mean of the score, almost all lecturers obtained a grade of B from the participants of the three batches. One instructor consistently got A, two instructors obtained A from one batch, and one instructor got C once.

Another form of the participants' evaluation toward the instructors was also given in the form of short remarks to answer the same questions. The list of the remarks which was confidentially given to each instructor includes positive comments as follows:

- diligent, kind, high sense of humor
- very attractive
- stimulating participants to think
- easily understood
- serving the participants sincerely

Polite criticisms as follows were also given:

- need to slow down a little
- good but cause some stress
- the methodology was sometimes not suitable
- did not master the methodology
- weak voice and weak classroom management

Facilities

Comments on learning facilities were addressed to the campus of IKIP Malang as the training site, classrooms, and particularly laboratories used for Listening Comprehension. In general, not many comments were made. There were more negative comments than positive ones.

The positive comments include:

- nice place, suitable for learning
- books were given freely
- facilities were sufficiently good
- good quality classrooms

The negative comments include:

- Sometimes the equipment for Listening did not work well
- Language Laboratory did not function, as it should have
- Library was not maximally utilized
- The TOEFL exercise books were not equipped with the cassettes
- Minimum teaching media
- No free access to the language laboratory.

The Grouping of the Participants

Little different from the responses to the other questions, the responses to the grouping aspect were all in the forms of complaints and suggestions. Some examples of their responses are as follows:

- The participants of different level of English competence should have been mixed.
- Because the participants of different fields were mixed together, the class was sometimes boring for the participants from English Departments.
- Placement of participants (should have considered) the different abilities.
- The grouping was not proportional. Many participants who did not know English were included in Group A (the upper group).
- The materials, the participants' competence, and the grouping were not suitable.

The comments above indicate that grouping would be the most difficult to handle because the groups were heterogeneous in terms of both their English proficiency and their fields of study.

CONCLUSION

Based on the data and the results of the data analysis, the following conclusions can be drawn.

- The minimum target of the training was Advanced competence and the highest achievement was reached by the participants who had at least Pre-Advanced competence at the beginning of the training.
- The recruitment of participants who minimally had Pre-Advanced competence would assure better achievement at the end of the program.
- Generally, participants from English and Natural Science Departments had higher initial English competence than participants from Social Sciences and Language (other than English) departments.
- Although generally participants from English Departments had high English competence, there were many of them from certain LPTKs whose English competence was not sufficient for being English lecturers. Many of them only had Pre-Advanced and even Intermediate level of English competence.
- The order of the participants' initial English competence related to their field background was also reflected in the level of achievement as shown in the gain achieved at the end of the training program.
- Feedback obtained from many ways, particularly from Final Program Evaluation revealed that the curriculum used as the base of the training was already considered solid.
- Although the main objective of the training is to increase the participants' TOEFL scores, at least for many participants the curriculum of the training used in PPS IKIP MALANG was considered more beneficial because it was TOEFL-Plus. In TOEFL-Plus, some activities with broader objectives, such as Speaking, Academic and Social Orientation, and even Seminar Practice were also provided.
- With some changes and improvements here and there, the teaching learning process would be considered to have an established pattern.
- Instructors were one of the important factors determining the success of the training. It was revealed that they had excellent academic qualification, quality, enthusiasm, dedication and teaching management.
- Many learning facilities need more attention and utilization, particularly the equipment for Listening Class exercises, such as the language laboratory and the tape recorders.
- From the correlation coefficients (0.52 - 0.73) between the International TOEFL and the TOEFL-Equivalent test developed by the program, it

is found that the TOEFL-Equivalent test has sufficiently high validity with the International TOEFL as the criterion, and therefore, it has the power of predicting the participants' scores in the International TOEFL.

RECOMMENDATION

Based on the above conclusions, the following are several recommendations.

- For the efficiency of the training and to achieve optimum result which will eventually satisfy all parties, participants should be required to have, at least, Pre-Advanced competence at the beginning of the program.
- Academic selection should be administered to obtain participants who meet the requirement recommended above.
- If the requirement on the minimal competence cannot be fulfilled for some reasons, Intermediate competence should be determined as the limit of the participants' competence level to be able to attend the training.
- If the requirement on the Pre-Advanced competence level cannot be met, the participants should be grouped into smaller groups, consisting of 10 to 12 participants to obtain optimal results.
- In addition to the smaller grouping, additional activities in the form of extra and/or tutorial classes with higher frequency and more individualized supervision should also be provided.
- Early preparation for International TOEFL can be made for participants with Advanced or higher competence. These participants should also be given the opportunity to take the International TOEFL as soon as they are ready without waiting at the end of the training program. Another activity which can be provided for these participants may include processing their plan to study overseas, selecting and determining the target university, filling out admission forms, and meeting other administrative requirements.
- The curriculum in the form of TOEFL-Plus as has been used so far should be retained.
- Native speakers involvement in the teaching activities should be retained and even increased by involving those from inside as well as from outside the campus.

- The equipment of language laboratory should be improved and its function should be optimized to support the learning activities, especially for Listening Comprehension classes and also for participants' independent study.

NOTE:

This article is based on the final report of the program evaluation conducted by the writer and M. Soenardi Djiwandono under the request of the Secondary School Teacher Development Project, Directorate General of Higher Education.

REFERENCES

- Brown, J.D. 1995. *The Elements of Language Curriculum*. Boston: Heinle & Heinle Publishers.
- Johnson, P. 1988. English Language Proficiency and Academic Performance of Undergraduate International Students. *TESOL Quarterly*, 22(1), 164—168.
- Lynch, B.K. 1990. A Context-Adaptive Model for Program Evaluation. *TESOL Quarterly*, 24(1), 23—42.
- Stufflebeam, D.L. and Shinkfield, A.J. 1985. *Systematic Evaluation*. Boston: Kluwer-Nijhoff Publishing.