

THE PROFILE OF EFL LEARNERS AS MEASURED BY AN ENGLISH PROFICIENCY TEST

Gunadi Harry Sulisty & Suharyadi
Universitas Negeri Malang, Indonesia
gunadi.hs@um.ac.id & suharyadi.fs@um.ac.id

Abstract: This study aims at profiling the students' English proficiency in the Department of English, *Universitas Negeri Malang* (UM). The English proficiency of 277 students across admission Classes in 2016, 2015, 2014, and 2013 who are involved in the study is measured by using a proficiency test. The subtests reliability was assured and the data was analyzed by using descriptive and inferential statistical analyses. The results show no respondent reached the advanced level; the majority was in the low intermediate level. It indicates that the respondents' score was not linear with their year of admission classes. However, a linearly increasing average score was respectively observed from the scores of Classes of 2015, 2014, and 2013. Statistical differences in the mean scores of the students across admission classes were observed. Next, no evidence of statistical differences was observed in the mean score of the students across different study programs in their seventh semester. Finally, none of Class of 2013 exactly met the graduate profile. However, those beyond the profile and those below the profile were marginally comparable. These findings were further discussed.

Keywords: *English Department students, language proficiency, Proficiency test*

INTRODUCTION

Categorized as one of the oldest English majors in Indonesia, the Department of English, Faculty of Letters, State University of

Malang (hereinafter referred to as the Department) promotes a vision to be a leading center of ICT-based higher education service provider that is superior and responsive to global change in the field of English language education and teaching as well as global developments in English language and literature (English Department Catalogue, 2014: 12 & 20). This global prospective has initiated the Department to set one of the main learning achievement targets of the graduates. It is focused on the competence in English equivalent to the scores of 525 points of ITP TOEFL/ of 196 points of CBT TOEFL/ of 69-70 points of iBT TOEFL/ of 5.5 points of IELTS' (English Department Catalogues, 2014 & 2016). This policy is established for the students' professionalism.

Hence, as an institution providing prospective workforces for educational and non-educational sectors, the expectation of both the government and the society in having proper education in the Department is quite high. The number of prospective students interested in studying in the Department is increasing while the seats available remain relatively the same. As a result, the admission ratio is constricted (Information Systems, State University of Malang, 2016). Due to such a selection tightness ratio, the chances for the Department to admit qualified students are widely open. With such opportunities, the learning atmosphere can be run in a conducive and effective way so that the outputs of the Department are qualified graduates.

However, thus far, no evaluation measure of the programs at the Department as the institution of education providers has been conducted systematically (an informal interview with the Head and the Secretary to the Head of the Department, January 2017). As a result, the effectiveness of learning in the Department has never been known empirically although in the past the Department once earned the recognition as the best English majors in ASEAN as an English learning institution. Based on this reality, one thing that remains interesting to examine with regard to the students' English proficiency is whether classroom learning affects learning outcomes as measured by the TOEFL as a learning profile of the Department.

Therefore, this research analyzes the language proficiency as one of kernels for professionally competent graduates.

Several researchers both from within and outside Indonesia have attempted to raise the issue pertaining to the relationship between learning activities and learning outcomes as measured by standardized tests. In the country, for example, there was evidence of impacts of learning on TOEFL training with learning outcomes as measured by TOEFL (Rakhmadi and Nurrohmah, 2009; Tjetje and Wulaningrum, 2012; Dewi, Darna, and Suprato, 2015; and Sucahyo, 2016). Similar research overseas has also been conducted. Through experimental research (Afsahi & Biria, 2016; Rahimirad, 2014), the research results revealed that experimental classes with the TOEFL materials demonstrated significant score differences. Robb and Ercanbrack (1999) disclose the impact of TOEIC test preparation training on the TOEIC scores on participants of English-language student group training and non-English study group students. Ling, Powers, and Adler (2014) show a moderate to high score on the TOEFL iBT tests with diverse learning achievement.

Correlational research between language proficiency measured by using the TOEFL and academic success as measured by using the grade point average (GPA) has also been performed. The research findings reveal an inconsistent relationship (Light, Xu and Mossop, 1987). On the other hand, Johnson (1988), Light and Teh-Yuan (1991), Hu (1991), Wait and Gressel (2009), and Stoyhoff (1997) revealed significant relationships. However, Cho, & Bridgeman (2012) found a weak link between the TOEFL iBT score and the GPA. Lo (2002), Krausz, Schiff, Schiff, & van Hise (2005), and Nelson, Nelson, & Malone (2004) concluded that the TOEFL cannot be used as a key predictor of predicting academic success. These studies in a way are theoretically based. Dunkin and Biddle (1974) state that the learning-product variables, in this case English proficiency, relate to process variables, presage variables and context variables.

The above-mentioned studies are deliberately designed using the TOEFL materials during the teaching learning processes. As far as

the present researchers' best knowledge, no research has been carried out to examine a link between EFL major learning and its learning outcomes as measured by the TOEFL. The Department has determined the graduates' English proficiency profile by using the TOEFL score as a benchmark of the learning outcome profile. However, so far there has been no systematic effort to know how far the of the Department's graduate profile is achieved. Therefore, it is necessary to conduct research to reveal student learning outcomes whether it is in accordance with the target of learning competencies or not if the Department's commitment is to make a meaningful contribution to the development of the nation and state in the provision of professional teachers who are qualified according to the Department's graduates' profile.

With reference to the background described in the previous section, this study aims (a) evaluating in general the level of English proficiency of the Department's students of classes of 2016, 2015, 2014 and 2013 based on the equivalent TOEFL score; (b) evaluating whether there are differences in the mean score of the English proficiency of different cross-class students (Classes of 2016, 2015, 2014, and 2013) based on the equivalent TOEFL score; (c) evaluating the level of students' English proficiency per class (Classes of 2016, 2015, 2014, and 2013) based on a per sub test score on equivalent TOEFL; (d) analyzing whether there is any difference in the TOEFL score before attending the program with the TOEFL score after attending the program in the third and the fourth academic years; (e) analyzing whether there is any difference in English proficiency between 7th semester students of the English Language Teaching Program (ELT) and 7th semester students of the English Language and Literature Study Program based on the TOEFL equivalent score; (f) analyzing whether the English proficiency of the 7th semester students, namely the English Language Teaching Study Program and the English Language and Literature Study (ELL) Program, has fulfilled the graduate profile as stated in the Catalogue of the

Department 2016 (page 14) based on the equivalent TOEFL score and with its percentage.

LITERATURE REVIEW

English Language Proficiency

The definitions of language proficiency vary. Chomsky defines language proficiency in terms of competence and performance. The former is about one's language knowledge in his/her mind while performance is about one's actual use of language in communication (Bacarić & Mihaljević Djigunović, 2007). Brumfit (1984:543) defines language proficiency as, "the maximally effective operation of the language system so far acquired by the student". Carter and Nunan (2001) also content that proficiency refers to one's application skills in second language for achieving communication. Similarly, Richards, Platt and Platt (1992) think that language proficiency is about how well a person can reach the levels of skills in using the language. It is then concluded that language proficiency is a person's ability of using the language for communication in both modes: written and spoken.

In the context of English language teaching, Li, Chen, & Duanmu (2010) & Wardlow (1999) argue that language proficiency plays a crucial role for students in completing their studies in English-medium institutions, especially for those whose first language is not English. Therefore, those who are considered having good language proficiency indicates that they must have mastered the skills or competence of a language needed so that they can perform well in their communication through both modes. Stern (1991) states that being proficient in a language means that students should master the four language skills. Richards (1978) adds that language proficiency includes syntax, semantic, discourse, and sociolinguistic elements. Bachman and Palmer (1996; 2010) state that language proficiency comprises organizational knowledge, grammatical knowledge, textual knowledge, pragmatic knowledge, functional knowledge and socio-linguistic knowledge.

Test of English as a Foreign Language (TOEFL)

Test of English as a Foreign language (TOEFL) is one of the standardized tests used widely to measure students' English language proficiency. As a proficiency test, TOEFL has been taken by non-native speakers (students) who are willing to study abroad (commonly in English speaking countries). In fact, more than one hundred countries (particularly international educational institutions, domestic higher-learning institutions and non educational agencies) and thirty millions people from different countries have used TOEFL since its first establishment in early 1960s (Sulistyo, 2009); (ETS, 2018).

TOEFL has two types of formats: internet-based test (IBT) and paper-and-pencil (PBT) version. Both formats do not differ significantly in terms of content for reading passages and recorded conversations and skills measured. The skills tested in the IBT test comprise four sections: reading, listening, speaking and writing while the PBT test consists of three sections: listening comprehension, structure and written expression, and reading comprehension (ETS, 2018). In terms of purpose, both formats aim at measuring one's English proficiency whose native language is not English. The scores resulted from those tests are useful in sense that they can be used to know how well a person is able to communicate in English in the academic setting.

There has been a lot of research conducted on the use of TOEFL. Studies undertaken by Rakhmadi and Nurrohmah, 2009; Tjetje and Wulaningrum, 2012; Dewi, Darna, and Suprato, 2015; and Sucahyo, 2016 evidenced that TOEFL impacted learning outcomes. Experimental research carried out by Afsahi & Biria, 2016; Rahimirad, 2014 also indicated that experimental classes with the TOEFL materials demonstrated significant score differences. Robb and Ercanbrack (1999) disclose the impact of TOEIC test preparation training on the TOEIC scores on participants of English-language student group training and non-English study group students. Ling, Powers, and Adler (2014) show a moderate to high score on the TOEFL iBT tests with diverse learning achievement.

Correlational research between language proficiency measured by using the TOEFL and academic success as measured by using the grade point average (GPA) has also been performed. The research findings reveal an inconsistent relationship (Light, Xu and Mossop, 1987). On the other hand, Johnson (1988), Light and Teh-Yuan (1991), Hu (1991), Wait and Gressel (2009), and Stoyhoff (1997) revealed significant relationships. However, Cho, & Bridgeman (2012) found a weak link between the TOEFL iBT score and the GPA. Lo (2002), Krausz, Schiff, Schiff, & van Hise (2005), and Nelson, Nelson, & Malone (2004) concluded that the TOEFL cannot be used as a key predictor of predicting academic success. These studies in a way are theoretically based. Dunkin and Biddle (1974) state that the learning-product variables, in this case English proficiency, relate to process variables, presage variables and context variables.

METHOD

This survey study was cross sectional aimed to examine the phenomena that developed contemporarily (Kerlinger, 1973; Tavakoli, 2012) concerning the Department's EFL students learning English with regards to the TOEFL score. The target population of this study is all students of the Department, Faculty of Letters, State University of Malang registered formally in the academic year 2015/2016, that is 840 (eight hundred and forty) students who were students of Classes of 2013, 2014, 2015 and 2016. Samples were drawn from samples of both the English Language Teaching and the English Language and Literature Study Programs selected by one class at random for each class: 2014, 2015, and 2016. For Class of 2013, all students are involved as research samples. With such a sampling strategy, this present study involved as many as 277 respondents. This was about .33 or 1/3 of the target population of the study.

The instrument used for this study is Barron's Practice Exercises for the TOEFL, 8th edition (Sharpe, 2015) which includes listening comprehension, grammar and written expression, and reading comprehension sub-tests. Instrument tryout results show

high reliability (Listening Comprehension $r=.811$, Grammar and Written Expression $r=.806$ and Reading Comprehension $r=.896$). The results of the item analysis show that the Listening Comprehension, Grammar and Written Expression, and Reading Comprehension subtests have a moderate difficulty level ($p = .569, .486, \text{ and } .640$; respectively); *r-pbis* coefficients were adequate ($.312, .339, \text{ and } .398$ respectively); discrimination indices *D* are adequate ($.427, .461, \text{ and } .537$ respectively).

The data was collected in accordance with the standard procedures normally performed in the administration of the TOEFL paper-pencil version. First, the listening comprehension sub-test was administered; which was then followed with the grammar and written expression test, and finally the reading comprehension section with a total allocated time of about 3 hours for all subtests. In this research the data collection was performed in the class. The data was collected between October and November 2016.

The data was scored dichotomously using paper-pencil TOEFL scoring. This scoring is performed using ITEMAN Version 3.50. The raw scores of each subtest were then converted using the conversion table specified in the Educational Testing Services (ETS) (ETS, 2018). The TOEFL converted scores of each class was then analyzed descriptively for the mean, mode, SD, variance, minimum score, and the maximum score. The English proficiency level of the Department's students of Classes of 2016, 2015, 2014 and 2013 in combination was analyzed using the English proficiency category based on the TOEFL score. For the purpose, recoding procedures in the SPSS version 21 program were performed.

All tests meant to examine mean differences in the English proficiency require the formulation of their corresponding statistical hypotheses and the statistical assumption tests, i.e. normality and data homogeneity with Kolmogorov-Smirnov test and homogeneity with Levene's test. If the statistical data assumption test is met satisfactorily, the statistical hypothesis test used is parametric statistics. Conversely, if the statistical data assumption test is violated,

the statistical hypothesis test used is non parametric statistics. However, considering the robustness of ANOVA (Hair, Anderson, Tatham, & Black, 1995), both non-parametric statistical analyzes of Kruskal-Wallis and ANOVA were used to compare the mean student TOEFL scores.

To analyze the level of students' English proficiency per class i.e. each of Classes of 2016, 2015, 2014, and 2013, the criteria for conversion scores per sub test on TOEFL are developed. Then, recoding data to determine the level of proficiency per subtest together with the frequency of each level of the proficiency as well as the percentage was made. Finally, to find out if the students' English proficiency meets the graduates profile as stated in the Catalogue of the Department of 2016 Edition (page 14), together with its percentage, recoding procedures on SPSS version 21 were also performed.

FINDINGS

The results of descriptive analyses of the English proficiency level of the Department's students' data of all classes of 2016, 2015, 2014, and 2013 are summarized and presented in Table 1. It shows that the lowest English proficiency of the Department's students is shown by a 360 point score. This score is slightly above the lowest possible TOEFL equivalent score of 337. Meanwhile, the highest English proficiency of the Department's students achieved is shown by a score of 616.67 points. This score is below the lowest possible TOEFL equivalent score of 626. This means that none of the respondents attained the 'advanced' category in the English proficiency level.

Table 1 Results of Descriptive Statistical Analyses of the Data of Classes of 2016, 2015, 2014, and 2013

Parameter	Statistics
Mean	*502.5030
Mode	*516.67

Parameter	Statistics
SD	*47.77995
Variance	*2282.923
Minimum score	*360.00
Maximum score	*616.67
N	277

*Original scores as computed

The English proficiency mean score of all the respondents is 502.5030, or rounded up to 503 points. Meanwhile, the mode is 516.67, or rounded up to 517 points. These figures show that in general the English proficiency of the respondents is categorized as low intermediate. Meanwhile, data based on its SD, that is 47.77995, or rounded up to 48 points, indicates that the respondents' English proficiency score are varied. Figure 1 illustrates the score distribution of the respondent's data.

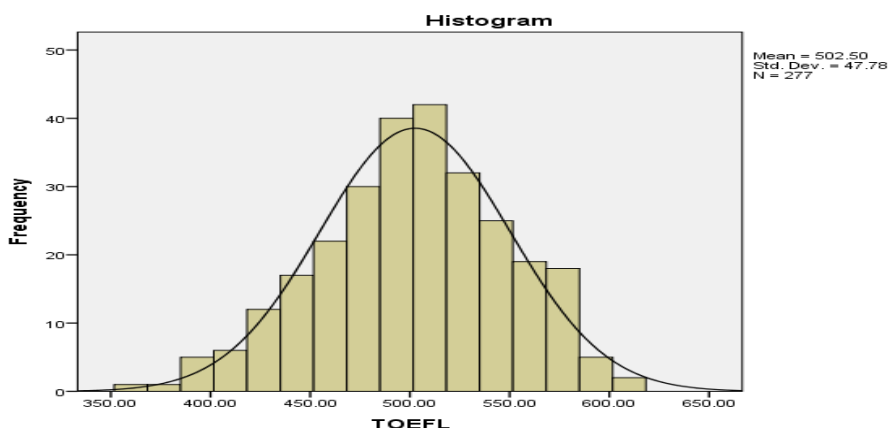


Figure 1 Score Distribution of All Respondents' Data on the TOEFL Score

The Department's students' English proficiency scores of all classes as shown in Figure 1 forms are nearly normally distributed. This distribution is close to the Gauss' theoretical distribution (Hair, Anderson, Tatham, & Black, 1995). The minimum score figures at 360 points while the maximum score reaches a figure of 616 points with a

mean score of 502 points. This means that the English proficiency of the Department's students of all classes ranges from the low to the upper intermediate levels of proficiency, although there are some who are in the advanced and the intermediate levels.

The results of the respondents' English proficiency profile analysis of per class' data are presented in Table 2.

Table 2 Descriptive Statistics of Per Class' Data: Classes of 2016, 2015, 2014 and 2013

Parameter	Class			
	2016	2015	2014	2013
Mean	505.6989	475.6614	483.9394	528.4192
Mode	516.67	436.67 ^a	443.33	526.67
SD	35.52565	49.75796	52.80249	35.19487
Variance	1262.072	2475.855	2788.103	1238.679
Minimum score	416.67	360.00	386.67	440.00
Maximum score	600.00	603.33	590.00	616.67
n	62	63	55	97

As shown in Table 2 the mean scores of Classes of 2016, 2015, 2014, 2013 are different, respectively figuring at 505,6989, 475,6614, 483,9394, and 528,4192, with the highest mean scores achieved by class of 2013, followed by class of 2016, 2014 and 2015. This trend is also reflected in the corresponding minimum score of 416.67; 360.00; 386.67, and 440.00. While the maximum score is also almost similar, except for Classes of 2015 and 2016. It is obvious based on their SD that the distribution of the scores in each class is also quite varied.

English Proficiency Level of All Classes of 2016, 2015, 2014, and 2013

Table 3 summarizes the results of the analysis on the level of English proficiency of the Department's students of all classes based on TOEFL equivalents.

Table 3 English Language Proficiency Level of Classes of 2016, 2015, 2014, and 2013

Level	Score range	f	%
Elementary	337 - 459	47	17.0
Low intermediate	460 - 542	171	61.7
High intermediate	543 - 626	59	21.3
Total		277	100.0

As shown in Table 3 the English proficiency level of the respondents of all classes is spread from 'elementary level' (337 - 459 points) to 'intermediate level' (543 - 626 points). No one reaches an 'advanced' proficiency level (627-677points). In addition, the distribution of students' proficiency level is also not the same. Most of the students (61.7%) are in the 'low intermediate' category while only a small number of students reach the category of 'elementary' (17%) and 'high intermediate' (21.3%).

Comparison of the Respondents' English Proficiency Levels

Is there any difference in the mean score in the English proficiency level across classes? This question seeks the answer whether there is a difference in the respondents' TOEFL average scores of classes of 2016, 2015, 2014, and 2013. The result of non-parametric statistical analysis - Kruskal-Wallis comparison of TOEFL score of students of classes 2016, 2015, 2014, and 2013 indicates significance .000 meaning that there is empirical evidence to reject the null hypothesis.

There is empirical evidence to accept the corresponding alternative hypothesis which states that there is a difference in mean TOEFL score of classes of 2016, 2015, 2014 and 2013. Further, taking into account the robustness of ANOVA, a summary of the results of ANOVA is presented in Table 4.

Table 4 Results of ANOVA of TOEFL Scores of Classes of 2016, 2015, 2014, and 2013

	Sum of Squares	df	Mean Square	F	Sig.
Between groups	130126.676	3	43375.559	23.685	.000
Within groups	499960.127	273	1831.356		
Total	630086.803	276			

The result of ANOVA shows the observed F value figuring at 23.685 with df. = 3. The F value in the critical table with a significance level of 0.05 is 2.62; meanwhile, the F value of in the critical table F with a significance level of 0.01 is 3.83. This shows that the observed F value > the F value in the critical table. The results of the analysis also showed a significance level <0.05 even 0.01, i.e. Sig. = .000. This means that there is empirical evidence to accept the alternative hypothesis stating that there is a difference in the mean score of the TOEFL student scores of classes of 2016, 2015, 2014, and 2013. This analysis supports the findings of the non-parametric - Kruskal-Wallis tests presented previously.

Considering the result of ANOVA on the difference of the mean scores described above, a post hoc analysis is conducted. This follow-up analysis is intended to examine which mean scores of Classes of 2016, 2015, 2014 and 2013 demonstrate differences in TOEFL mean scores. The results of the follow-up analysis are summarized and presented in Table 5.

Table 5 Results of Post-Hoc Analysis across Classes

Comparison between Classes	Mean Difference	SD	Significance.
2016	2015	30.03755*	7.65553 .000
	2014	21.75953*	7.92688 .006
	2013	-22.72032*	6.95830 .001
2015	2016	-30.03755*	7.65553 .000
	2014	-8.27802	7.89725 .295
	2013	-52.75787*	6.92453 .000

	2016	-21.75953*	7.92688	.006
2014	2015	8.27802	7.89725	.295
	2013	-44.47985*	7.22339	.000
	2016	22.72032*	6.95830	.001
2013	2015	52.75787*	6.92453	.000
	2014	44.47985*	7.22339	.000

As shown in Table 5, there is empirical evidence of statistical differences in the mean score across Classes. The mean score of Class of 2016 differs from the mean scores of Classes 2015, 2014 and 2013 at the significance level <0.01 . Secondly, the mean score of Class of 2015 differs from the mean score of Classes of 2016, and 2013 at the significant level <0.01 , but does not differ from the mean score of Class of 2014 (sig 0,295). Thirdly, the mean score of Class of 2014 differs from the mean score of Classes of 2016, and 2013 at the significance level $<0,01$, but does not differ from the mean score of Class of 2015 (sig 0,295). Fourth, the mean score of Class 2013 differ from the mean score of Classes of 2016, 2015 and 2014 at significance level <0.01 .

Level of English Proficiency across Classes on Sub-tests

The results of the English proficiency level analysis of students across Classes of 2016, 2015, 2014, and 2013 based on the sub-tests of listening comprehension, grammar and written expression, and reading comprehension are presented and summarized in Table 6.

Table 6 English Language Proficiency Level of Class of 2016

Level	Listening Comprehension		Grammar and Written Expression		Reading Comprehension	
	f	%	f	%	f	%
	Elementary	0	0	1	1.6	0
Low intermediate	0	0	34	54.8	3	4.8
High intermediate	45	72.6	26	41.9	50	80.6
Advanced	17	27.4	1	1.6	9	14.5
Total	62	100.0	62	100.0	62	100.0

The result of data analysis in Table 6 above informs that the level of English proficiency of Class of 2016 based on scores of listening comprehension, grammar and written expression, and reading comprehension differs. In general in terms of language skills, both listening comprehension (72.6% and 27.4%) and reading comprehension (80.6% and 14.5%), are mostly in the 'high intermediate' and 'advanced' levels. Meanwhile, on the mastery of the language component (grammar and written expression), the majority is in the 'low intermediate' level of mastery. Thus, the mastery of Class of 2016 on the language component - grammar and written expression - is still under the mastery of listening comprehension and reading comprehension.

Table 7 presents the results of the data analysis to determine the level of English proficiency of the Class of 2015 based on the sub-tests of listening comprehension, grammar and written expression, and reading comprehension.

Table 7 English Language Proficiency Level of Class of 2015

Level	Listening Comprehension		Grammar and Written Expression		Reading Comprehension	
	f	%	f	%	f	%
	Elementary	0	0	2	3.2	1
Low intermediate	7	11.1	35	55.6	16	25.4
High intermediate	53	84.1	25	39.7	36	57.1
Advanced	3	4.8	1	1.6	10	15.9
Total	63	100.0	63	100.0	63	100.0

A similar pattern in the proficiency level of Class of 2015 with that of Class 2016 is observed. Table 7 shows the differences in the level of English proficiency of Class of 2015 based on the scores of listening comprehension, grammar and written expression, and reading comprehension. In addition, Class of 2015 language skills, both listening comprehension (84.1% and 4.8%) and reading comprehension (57.1% and 15.9%), are in the majority of 'high intermediate' and 'upper' levels. Meanwhile, in terms of mastery of the language component (grammar and written expression), the

majority of Class of 2015 is in the 'low intermediate' mastery level (55.6% and 3.2%).

Table 8 presents the results of the data analysis to determine the level of English proficiency of the Class of 2014 based on the sub-tests of listening comprehension, grammar and written expression, and reading comprehension.

Table 8 English Language Proficiency Level of Class of 2014

Level	<i>Listening Comprehension</i>		<i>Grammar and Written Expression</i>		<i>Reading Comprehension</i>	
	f	%	f	%	f	%
	Elementary	0	0	1	1.8	1
Low intermediate	5	9.1	30	54.5	8	14.5
High intermediate	40	72.7	22	40.0	38	69.1
Advanced	10	18.2	2	3.6	8	14.5
Total	55	100.0	55	100.0	55	100.0

Table 8 shows similarities in the pattern of English language proficiency level of Class of 2014 with those of Class of 2015 as well as Class of 2016. As shown in Table 9, there is a difference in the level of English Language proficiency of Class 2014 students based on scores on listening comprehension, grammar and written expression, and reading comprehension test on TOEFL equivalent. In addition, the proficiency levels of Class of 2014 in both listening comprehension (72.7% and 18.2%) and reading comprehension (69.1% and 14.5%), are majority in the 'high intermediate' and 'advanced' levels. Meanwhile, the majority is in the mastery levels 'low intermediate' (54.5% and 1.8%) in the grammar and written expression.

Table 9 presents the results of the data analysis to determine the level of English proficiency of the Class of 2013 based on the sub-tests of listening comprehension, grammar and written expression, and reading comprehension.

Table 9 English Language Proficiency Level of Class of 2013

Level	<i>Listening Comprehension</i>		<i>Grammar and Written Expression</i>		<i>Reading Comprehension</i>	
	f	%	f	%	f	%
	Elementary	0	0	0	0	0
Low intermediate	4	4.1	6	6.2	2	2.1
High intermediate	79	81.4	70	72.2	65	67.0
Advanced	14	14.4	21	21.6	30	30.9
Total	97	100.0	97	100.0	97	100.0

Data in Table 9 shows that none of Class of 2013 reached an English proficiency level in the 'elementary' category in listening comprehension, grammar and writer expression, and reading comprehension subtests. A few are still in the 'low intermediate' category these three sub-tests (i.e. 4.1%, 6.2% and 2.1% respectively). Class of 2013 tended to be in the 'high intermediate' and 'advanced' mastery levels in the three sub-tests. However, their mastery on the language component - grammar and written expression - (72.2% and 21%) is lower than the mastery of listening comprehension (81.4% and 14.4%) and reading comprehension (67.0% and 30.9%).

Comparison of the TOEFL Mean Score Before and After Class Attendance

Question 5 is formulated as follows: Is there any difference in the initial TOEFL mean score prior to students attending classes with that TOEFL mean score after attending lectures in the third and fourth year? The TOEFL score data for the 4th year in this study is the data taken from the TOEFL score of Class of 2013 students. The data of Class of 2013 was taken before the lecture began formally, i.e. before the first semester began in 2013 and the data when they were in the 7th semester by 2016. The results of the t-test analysis are summarized in Table 10.

Table 10 The Result of t-test Analysis for the TOEFL Mean Scores Before and After Class Attendance of Class of 2013

Score	t	df	Sig. (2-tailed)	Mean	95% Confidence Difference Interval	
					Lower bound	Upper Bound
By 7 th Semester	147.872	96	.000	528.41924	521.3259	535.5126
Before Class Attendance	98.659	96	.000	421.31959	412.8428	429.7964

The results of t test analysis presented in Table 10 show a statistically significant difference (Sig. 2-tailed = 0,000). This means that the mean score of Class of 2013 by the 7th semester is different from that at the beginning before the lecture began.

To provide further evidence, mean score data of Class of 2014 in this study were taken from the TOEFL score in the 3rd year. Like the data of the Grade 2013 students, the data of the Class 2014 students were taken at the beginning before the lecture begins, i.e. before the first semester began in 2014 and the data when they were in semester V in 2016. The results of comparing the mean score are summarized in Table 11.

Table 11 The Result of t test Analysis for the TOEFL Mean Scores before after Class Attendance of Class of 2014

Score	t	df	Sig. (2-tailed)	Mean	95% Confidence Difference Interval	
					Lower bound	Upper Bound
By 7 th Semester	67.970	54	.000	483.93939	469.6649	498.2139
Before Class Attendance	58.220	54	.000	420.69091	406.2039	435.1779

Further evidence indicates that the results of t test analysis presented in Table 11 show a statistically significant difference (Sig. 2-tailed = 0,000). This means that the mean score of Class of 2014 by the 7th semester is different from that at the beginning before the lecture began.

Comparison of the TOEFL Mean Scores across Study Programs

Data analysis to find out whether there is evidence of a statistical difference of in the mean score the 7th semester student of English Language Teaching Study Program and that of the 7th semester student of English Language and Literature Study Program based on TOEFL score is performed by using t-test. The results of comparing the mean score are summarized in Table 12. The results of the t-test analysis show the significance value of t test of 0.603 for the score with the same variance assumption and 0.619 for the score with unequal variance assumption. Both of these values are much larger than 0.05 (Sig. 2-tail).

This means there is not enough empirical evidence to reject the null hypothesis. The implication of acceptance of this null hypothesis is to reject alternative hypotheses. Thus, there is no empirical evidence to state the difference in mean score of the TOEFL mean score of 7th semester students of English Language Teaching Study Program and that of 7th semester students of English Language and Literature Study Program. In other words, empirically the mean score in the TOEFL of 7th semester students of English Language Teaching Study Program and that of 7th semester students of English Language and Literature Study Program do not differ. This can be analyzed from the statistical result conducted in this research. The detailed analysis of two study programs of the 7th semester students is presented in Table 12.

Table 12 The Result of t test Analysis for the TOEFL Mean Scores of the 7th Semester Students of English Language Teaching Study Program and that of the 7th Semester Student of English Language and Literature Study Program

Score	t	df	Sig. (2-tailed)	Difference	95% Confidence Difference Interval	
					Lower bound	Upper Bound
Equal variances assumed	0,521	97	0,603	3,83808	7,36235	-10,77416
Equal variances not assumed	0,500	68,075	0,619	3,83808	7,68116	-11,48914

Achievement of the Departments Students Profile Based on the TOEFL Score

The 7th semester student, that is Class of 2013 of both English Language Teaching Study Program and English Language and Literature Study Program, are expected to achieve the Department profile by the end of their study. To answer that question, 97 respondents of 7th semester students, or Class of 2013, were drawn as the research sample. The results of the descriptive statistical analysis of their data are presented in Table 13.

Table 13 Achievement of the Departments Students Profile Based on the TOEFL Score

Category	Score	f	%
Under the Profile	337 - 524	43	44.3
Reaching the Profile	525	0	0
Beyond the Profile	526 - 677	54	55.7
Total		97	100.0

The results of data analysis as shown in Table 13 reveal that none of the English proficiency of 7th semester students, namely Class 2013, both English Language Teaching Study Program and

English Language and Literature Study Program reaches the profile of graduates as set forth in Catalogue of Department 2016 (page 14). However, half of the respondents (54%) have the English proficiency beyond the graduate profile, while the rest (43%) are under the profile.

DISCUSSION

An interesting finding from this study is that there is a tendency for lower ability of respondents in the language component revealed through the grammar and written expression sub-test than that in the listening and reading comprehension sub-tests. The findings of this study contradict the findings of Dewi, et al., (2015) and Suchyo (2016) studies which reveal that listening comprehension skills - their study respondents are a major problem. Conceptually, language elements should support the performance of language skills.

The lower ability in the language components in this study can be explained from two aspects. Firstly, in the Department learning English, including English grammar, no longer uses an approach that merely emphasizes the accuracy of language forms, but more on the meaningfulness in the context of the use of contextual communicative language forms. Learning in the English grammar course does not address discrete elements of grammar, but on learning that is oriented towards meaningful classroom activities rather than solely on the accuracy of language forms. As a result, the mastery of the meaning of language-based forms becomes more dominant. The grammar and written expression sub-test in the paper-and-pencil TOEFL is more likely to test aspects of grammatical sensitivity (Sulistyo, 2001). With such a form of testing, it is argued that the respondents of this study were also not familiar with the format of error recognition as in the equivalent TOEFL paper and pencil format.

Secondly, the habit of using a particular test format in some cases relates to performance on the test. Conceptually, some experts such as Bachman (1990), Alderson, Clapham, and Wall (1995),

Bachman and Palmer (1996), and Weir (2005) have suggested a link between the test format and the performance on the test. Similarly, empirically the relationship between the test format and the performance on the test ever performed. The Goldberg and Pedulla (2002) study revealed that the use of multiple choice formats performs better than the use of computer-based tests. The results of research conducted by Kobayashi (2002) revealed that the reading organization and test format had a significant influence on the performance of the respondents on the comprehension reading test. Salehi and Sanjareh (2013) also found similarities in conventional multiple-choice and multiple-choice tests modified on grammaticality judgment tests.

Students of the Department, Faculty of Letters, State University of Malang are not specially trained for the TOEFL, but they learn English to achieve certain competences specified by the Department with a graduate profile that has language skills equivalent to the TOEFL ITP with score 525 (Catalogue 2016 Edition, 2016 : 14). While TOEFL is more academically oriented in relation to the assumptions of English language proficiency for further study skills in the instructional context (Jamieson, Jones, Kirsch, Mosenthal, and Taylor, 2000: 1), the orientation of learning English in the Departments is different, controlled by competences as described in the Catalogue of 2016 Edition. This shows there is a difference in achievement orientation. With the aim of achieving the competences specified in Department's catalogue, the determination of this profile by the Department is considered quite realistic. Thus, the finding that none of the respondents of this study achieved English proficiency based on the equivalent TOEFL scores in the 'advanced' category (627-677 points) can also be seen as realistic.

Additionally, this study revealed that there is fair distribution of scores approaching the theory. With a mean score of 503 points, the distribution of student scores of Classes of 2013, 2014, 2015, and 2016 formed a spread with a normal bell shape (Hair, et al., 1995: 34). That is, the majority of students are in low intermediate category skills but there are respondents who are in the area of proficiency in the

elementary category (337-459 points). Besides, some are in the category of high intermediate (543-626 points).

The findings of this study are similar to the findings of Sucahyo's (2016) study which revealed that none of his research respondents achieved language proficiency at an advanced level. However, in contrast to the findings of Sucahyo's (2016) study which revealed the mean TOEFL score of respondents was 397 points. With the distribution of students' majority proficiency levels in the 'elementary' category, some at low intermediate levels, and only a small percentage of the intermediate level. The present study revealed that the mean score was 503 points, greater than the respondents' score on the Sucahyo (2016) study.

Important information derived from the results of this research analysis revealed that the mean score of TOEFL respondents of Classes of 2014 and Class 2013 taken after the lecture differed significantly statistically from the TOEFL score at the beginning before the lecture began. The findings of this study are consistent with the results of Tjetje and Wulaningrum (2012) studies which also reveal a difference between the mean score of the TOEFL students of the first semester and the final semester students although the level of proficiency in their study was classified as 'low intermediate'. Similar to the research findings of Tjetje and Wulaningrum (2012), the findings of this present study are also in line with the findings of Ling, et al., (2014) research. They revealed that subjects who study English in training programs can improve their language skills at least to moderate levels over the course of 6 months or more as indicated by the change in the TOEFL iBT practice test scores. The findings of this study are also in line with research findings Dewi, Darna, and Suprato (2015) which revealed that the TOEFL score of TOEFL trainees showed a significant improvement at the end of the training period compared to the initial score.

Although not achieving the TOEFL score in the 'advanced' category, in this study the respondents TOEFL score increases and other relevant studies mentioned above indicate the significance or

the contribution of the Department learning activities to the increase of the TOEFL score. Although there is no statistically significant difference between the TOEFL scores between Classes of 2014 and Class 2015, the empirical findings indicate that the mean the TOEFL scores from Classes of 2015, 2014 and 2013 consistently increase, i.e., respectively as follows: 475,661, 448,3939, and 452,84,192 points. Even if there is an anomaly occurring in Class of 2016, this phenomenon may be referred to as a case. Referring to the Dunkin and Biddle conceptions (1974), it is likely that Class of 2016 input is better than that of the previous classes, Classes of 2015, 2014, and 2013, for example. The results are informative of instructional functionality. If this is true, and the learning process can improve student competence, it is likely that in the fourth year, the equivalent TOEFL score in the advanced category will be achieved by Class of 2016.

The role of learning in improving English proficiency is in line with the concept of Dunkin and Biddle (1974) that the learning-product variables, in this case English proficiency, relate to other prior variables, mainly process variables, in addition to presage variables and context variables (Dunkin and Biddle, 1974: 38-43). Presage variables are, among others, related to lecturer attributes; meanwhile context variables are closely related to class context and student attributes.

CONCLUSION

Several points can be drawn as conclusions. These, however, need to be interpreted carefully based on facts and research findings pertaining to research problems. The graduate profile of the Department set up at the TOEFL ITP score of 525 points can be achieved empirically although not all candidates have achieved the profile yet. A comparison between those who have exceeded the benchmark score of 525 points and those still below the benchmark score indicates some points of considerations that need attention. This is understandable as when this study was conducted, the majority of

the respondents did not yet fully complete their studies in the Department.

This study does not claim as a final and comprehensive finding. There are several weaknesses that limit the generalization of the findings and results of this study. Firstly, the instruments used in this study are not the actual TOEFL tools although the instrument reliability used in this study is considered adequate from the reliability aspect (.70s). The aspect tested may not have fully measured the respondents' English proficiency as a standard TOEFL test. Therefore, the scores of measurements obtained in this study need to be interpreted with caution.

The aspects of writing and speaking are beyond the scope of this research. English proficiency in the present research is understood in terms of listening comprehension, grammar and written expression, and reading comprehension as measured by the TOEFL equivalent only. The present research does not include writing and speaking skills, while there are other procedures that can be used to indicate college students' language proficiency. In addition, not all students were involved as respondents in this study. In this present study, only half of Classes of 2016, 2015, and 2014, and most of the Class 2013 students drawn randomly as respondents of this present study. The skills of the students of the Department of in this study are thus limited only to the students involved as respondents of this study, excluding students who were not involved as respondents to this study.

Not all data collection processes of this research were performed with the administration standards of the real TOEFL test. As data collection is dependent on students' availability, some data collection processes were administered indirectly for 3 (three) consecutive hours.

With regards to the findings that not all the 7th semester students achieved the minimum standard, the learning and instruction on the courses offered by the Department needs to be optimized further. Optimization of learning for example can be made

possible by updating and enriching learning materials, making innovations in the teaching practices, and enriching the source of students learning. With the development of information technology, activities to update and enrich the learning materials and enlarge the learning resources can be optimized. The use of internet, for example through relevant independent assignments can be implemented.

Future researchers are recommended to conduct similar research using the TOEFL of different test modes: paper-and-pencil and internet-based to see the comparability of the score. In addition, studies may also be conducted using other instruments, such as IELTS, or studies that compare the scores obtained through the TOEFL and IELTS tests.

This study covers only aspects of the skills that are still global. Therefore, it is necessary to measure all aspects of language skills oriented to the diagnosis of the advantages and disadvantages of students as measured using the TOEFL and IELTS tests, or any others. In addition, this study has not revealed what factors contribute to the TOEFL or other scores as well as areas of weaknesses and strengths in learning English. Therefore, researchers may also undertake exploratory research that attempts to reveal the factors contributing to the language proficiency score and the areas of weaknesses and strengths.

REFERENCES

- Afsahi, S.E. & Biria, R. (2016). The impact of student-based instruction on improving IBT TOEFL scores of Iranian students. *International Journal of English Linguistics*. 6(7), 99-104.
- Bacarić, V. and Mihaljević Djigunović, J. 2007. Defining Communicative Competence. *Metodika: Journal of Theory and Application of Teaching methodologies in Preschool, Primary, Secondary and Higher Education* 8 (14), 94-103.
http://hrcak.srce.hr/index.php?show=clanak&id_clanak_jezik=42651

Sulistyo & Suharyadi, The Profile of EFL Learners as Measured by an English Proficiency Test

- Bachman, L. F., & Palmer, A. S. (1996). *Language testing in practice*. Oxford: Oxford University Press.
- Bachman, L. F., & Palmer, A. S. (2010). *Language testing in practice*. Oxford: Oxford University Press.
- Brumfit, C. (1984). *Communicative methodology in language teaching*. Cambridge: Cambridge University Press.
- Carter, R., & Nunan, D. (2001). *The Cambridge guide to teaching English to speakers of other languages*. (4th ed.). Cambridge: Cambridge University Press.
<http://dx.doi.org/10.1017/CBO9780511667206.013>
- Cho, Y. & Bridgeman, B. (2012). Relationship of TOEFL iBT[R] scores to academic performance: some evidence from American Universities. *Language Testing*, 29(3), 421-442.
- Dewi, I.I., Darna, D., & Suprato, D. 2015. The changes of students' TOEFL score after one year learning. *Humaniora*, 6(4), 507-519.
- Dunkin, M. J., & Biddle, B. J. (1974). *The study of Teaching*. Washington, DC: University Press of America.
- Educational Testing Service. (2001). *TSE and SPEAK Score User Guide*. Princeton, NJ: Educational Testing Service
- Educational Testing Service. (2018). *The TOEFL Test*. Accessed from <https://www.ets.org/toefl>.
- English Department Catalogue. (2014). Malang: Fakultas Sastra, Universitas Negeri Malang
- English Department Catalogue. (2016). Malang: Fakultas Sastra, Universitas Negeri Malang

- Goldberg, A.L. & Pedulla, J.J. (2002). Performance differences according to test mode and computer familiarity on a Practice GRE. *Journal of Educational and Psychological Measurement*, 62(6), 1053-1067.
- Hair, J.F., Anderson, R.E., Tatham, R.L., & Black, W.C. (1995). *Multivariate Data Analysis with Readings*. 4th Edition. London: Prentice-Hall International, Inc.
- Hu, S.C.P. (1991). English proficiency and academic performance of international graduate students. *Dissertation Abstracts International*, 52, 1626A
- Informal interview with the Head and the Secretary to the Head of the Department, Faculty of Letters, Universitas Negeri Malang, January 2017.
- Information System Universitas Negeri Malang*. (2016). *Dokumen Rasio Keketatan Seleksi Calon Mahasiswa Universitas Negeri Malang*.
- Jamieson, J., Jones, S., Kirsch, I., Mosenthal, P., & Taylor, C. (2000). *TOEFL 2000 Framework: A Working Paper*. Princeton, N.J.: Educational Testing Service
- Johnson, P. (1988). English language proficiency and academic performance of undergraduate international students. *TESOL Quarterly*, 22(1), 164-168.
- Kerlinger, F.N. 1973. *Foundations of Behavioral Research 2nd Edition*. New York: Holt Rinehart and Winston
- Kobayashi, M. (2002). Method effects on reading comprehension test performance: text organization and response format. *Language Testing*, 19 (2), 193-220.
- Krausz, J.S., A, Schiff, J, & van Hise, J. (2005). The impact of TOEFL scores on placement and performance of international students in the initial graduate accounting class. *Accounting Education*, 14(1), 103-111.

- Li, G., Chen, W., & Duanmu, J-L. (2010). Determinants of international students' academic performance: A comparison between Chinese and other international students. *Journal of Studies in International Education*, 14(-), 389-405. doi:10.1177/1028315309331490
- Light, R. L., Xu, M. & Mossop, J. (1987). 'English proficiency and academic performance of international students'. *TESOL Quarterly*, 21(2), 251-261.
- Ling, G., Powers, D. E., & Adler, R. M. (2014). Do TOEFL iBT® Scores Reflect Improvement in English-Language Proficiency? Extending the TOEFL iBT Validity Argument: Do TOEFLiBT® Scores Reflect Improvement in English-Language Proficiency? *ETS Research Report Series*, 2014(1), 1-16. <https://doi.org/10.1002/ets2.12007>. Assesed from <https://files.eric.ed.gov/fulltext/EJ1109334.pdf>.
- Lo, J.W. (2002). The relationship between TOEFL scores and first year GPA: A Study of freshmen international students attending Texas A & M University-Kingsville from 1996-2001. *Dissertation Abstracts International*, 63, 4193A-4194A.
- Nelson, C.V., Nelson, J.S., & Malone, B.G. 2004. Predicting success of international graduate students in an American university. *College and University*, 80(1), 19-27.
- Rahimirad. M. 2014. The Impact of Metacognitive Strategy Instruction on the Listening Performance of University Students. Proceedings pada International Conference on Current Trends in ELT, *Procedia - Social and Behavioral Sciences*, 98(2014), 1485 - 1491. doi: 10.1016/j.sbspro.2014.03.569 accessed from <http://www.sciencedirect.com/science/article/pii> on December 25, 2016.

- Rakhmadi, A. & Nurrohmah, R. 2009. *Evaluasi pengaruh pembelajaran berbasis Web untuk Bahasa Inggris sesi Structure and Written Expression*. Seminar Nasional Informatika (SEMNASIF), UPN Veteran Yogyakarta, 23 Mei 2009. Accessed from <http://jurnal.upnyk.ac.id/index.php/semnasif/article/view/929> on December 27, 2016.
- Richards, J. C., Platt, J., & Platt. H. (1992). *Dictionary of language teaching and applied linguistics (2nd Ed.)*. London: Longman.
- Richards, J. C. (1978). *Models of language use and language learning*. Rowley, Mass: Newbury House.
- Robb, T.N. & Ercanbrack, J. (1999). A study of the effect of direct test preparation on the TOEIC scores of Japanese university students. *TESL-EJ*. Volume 3(4), Accessed from <http://www.tesl-ej.org/wordpress/issues/volume3/ej12/ej12a2/> on December 16, 2016.
- Salehi, M. & Sanjareh, H.B. (2003). The impact of response format on learners' test performance of grammaticality judgment tests. *Journal of Basic and Applied Scientific Research*. 3(2), 1335-1345.
- Sharpe, P.J. (2015). *Barron's Practice Exercises for the TOEFL*, 8th Edition. New York: Barron's Educational Series, Inc.
- Stern, H. H. (1991). *Fundamental concepts of language teaching*. Hong Kong: Oxford University Press.
- Stoyhoff, S. (1997). Factors associated with international students' academic achievement. *Journal of Instructional Psychology*, 24(1), 56-68.
- Sucahyo, S.A. (2016). Peta nilai TOEFL mahasiswa Program Studi Pendidikan Bahasa Inggris Institut Agama Islam Negeri Samarinda. *Fenomena*, Volume 8(1), 101 - 109.

Sulistyo & Suharyadi, The Profile of EFL Learners as Measured by an English Proficiency Test

- Sulistyo, G.H. 2001. Several Technical Considerations for Taking the (Paper-and-Pencil-Based) TOEFL Test. *TEFLIN Journal*, 12(2), 223-241.
- Sulistyo, G.H. 2009. TOEFL in a Brief Historical Overview from PBT to IBT. *Bahasa dan Seni Journal*, 37(2), 116-127.
- Tavakoli, H. 2012. *A Dictionary of Research Methodology and Statistics in Applied Linguistics*. Tehran: Rahnama Press.
- Tjetje, N.F. & Wulaningrum, R. 2012. Hubungan antara lama waktu pembelajaran dan kemampuan Bahasa Inggris Mahasiswa Jurusan Akuntansi Politeknik Negeri Samarinda. *Jurnal Eksis*, 8(2), 2168 - 2357.
- Wait, I.W. & Gressel, J.W. 2009. Relationship between TOEFL score and academic success for international engineering students. *Journal of Engineering Education*, 98(4), 389-398.
- Wardlow, G. 1999. International students of Agriculture in U.S. institutions precursors to academic success. *Journal of Agricultural Education*, 30(-), 17-22.
- Weir, C. J. (2005) *Language Testing and Validation: An Evidence-Based Approach*, Basingstoke: Palgrave Macmillan.
- Wendt, A. & Woo, A. (2009). *A Minimum English Proficiency Standard for the Test of English as a Foreign Language Internet-Based Test (TOEFL iBT)*. Chicago, IL.: National Council of State Boards of Nursing, Inc. (NCSBN).