

# J E T L I

## Journal of Language Teaching and Linguistics

<https://ejournal.unida-aceh.ac.id/index.php/jetli>

### GENDER ISSUE IN TOEFL TEST: DOES IT MAKE SIGNIFICANT DIFFERENCE?

<sup>1\*</sup> Ida Muliawati

<sup>1</sup> Department of English Education, Faculty of Teacher Training and Education, Iskandarmuda University, Banda Aceh, Indonesia

\*Corresponding author: [ida.thesyarils@gmail.com](mailto:ida.thesyarils@gmail.com)

#### Abstract

Gender issue has been noted as a debatable issue over decades in second language learning, including in TOEFL test which is recognized throughout the world as a certified foreign language test. The focus of this paper is to explore and to compare the TOEFL proficiency between male and female students. The samples of this research were 20 students of Syiah Kuala University; 10 male students and 10 female students using simple random sampling. A descriptive quantitative research design was manipulated to sum up the data and the instrument deployed was TOEFL tests. Result of the tests shows that gender issue does not influence the TOEFL proficiency. It means that there is no significant difference in TOEFL test score between male and female students.

**Keywords:** *Gender, TOEFL, Language Test, Proficiency Test*

#### 1. Introduction

It is undeniable that the English proficiency is very essential in every aspect of life. One of the English proficiency tests which has been noted crucial and widely acceptable across the globe is TOEFL. Test of English as a foreign language is has been tremendously recognized as a standardized language testing throughout the world (Warfield et al., 2013, p. 196). mentioned that standardized test is defined as a test in which equal procedures and requirements are applied with a consistent standard and structured format such as time limits, response format, number of questions, and scoring format (Popham, 1999., (Brown, 2003).

In Indonesia, the TOEFL certified internationally is hold by the English Language System (ETS) (Mahmud, 2014). The university may also administer the TOEFL-like or TOEFL equivalent test, or TOEFL prediction, such as by center of language studies or by language laboratory within the university. Syiah Kuala University as the largest and the oldest national university in Aceh, Indonesia also

demands their students to take TOEFL test before their final exams with the score of 475. If the students cannot achieve the determined score they can not apply for the thesis examination.

Having such a good TOEFL score is not an easy task to do, since Most of EFL learners find it is hard to engage with English. To be precise, the TOEFL solely does not test just one skill but it also tests three skills altogether, listening, structure and written expression, and reading skills which is not easy at all for EFL learners since their first language is not English (Mahmud, 2004). It can be said that if one EFL learners have never get in touch the language in their life, then when they enroll in the test they will find it difficult to comprehend. Munadia (2016) in her paper mentions that students faced a lot of difficulties in mastering TOEFL test.

There are some factors which influence students as the TOEFL test-takers to success in the test—test-taking behaviors. Abundant researches related to the terms mentions that test-taking behaviors such risk-taking strategies, test motivation, and level of anxiety is important for test performance beside content knowledge (Stenlund et al. 2017). Similarly, some previous studies conducted by Dodeen et al . (2014); Naylor (1997), Stenberg (1998) revealed that it is important to minimize the degree of anxiety as well as tighten up the level of motivation in order to be successful in test. Bicak (2013) adds up that another factor to be a successful test-taker is that the willingness to take a risk in test. Furthermore, it is interesting to note that based on the research undergone by Stendlund et al (2017); Ellis & Ryan (2003); Hong et al. (2006) mention that male students tend to be braver in taking risk while answering test items compared to female students who experience high degree of anxiety while facing the test. Baldiga (2014) claims that female students tended to skip questions compared to male students who point out to eagerly guess the items. On the other hand, study from Stenlund et al. (2016) shows in contrast; female students recently reported to keen on guessing items when they are encountering with the difficult questions.

Moreover, gender comes as another variable researched in this study. Lee (2018) mentions that researches in determining whether gender influence TOEFL scores among female and male learners is still insufficient. Hyde & Linn (1998) conducted a research to seek out the significant differences in language performance between male and female students. They ended up with the conclusion that there is no gender difference in between them in absorbing English language. In a linear fashion, Lin & Wu (2003) also claims that gender does not determine the language capability. On the other hand, Cole (1997) found out that female students are more advantageous compared to males. In some language skills such as reading and writing, females reveal brighter than males.

Since the outcomes are still debatable, the researcher intends to conduct a new research underlying this issue. Hence, this paper examines and compares the toefl proficiency, especially in toefl PBT scores between female and male students in Syiah Kuala University.

## 2. Literature Review

Proficiency test determines and measures a learner's level of language. The examples of proficiency test are TOEFL and IELTS. Both of them are truly well-known among students who decide to pursue a scholarship or study in English-speaking universities (Nurhayati & Gir, 2014; Sugeng et al. 2012). Some countries regulate TOEFL and IELTS as one of the compulsory requirement to apply for their universities. Those two certified tests are needed to examine the students' proficiency and comprehension on English language, so they immensely portray or represent their ability in it. In other words, since they are language testing, they are strongly connected to test takers' abilities in understanding a certain language area in a certain test circumstance.

TOEFL, in particular, is standardized high-stake test to assess, investigate, examine, measure, and sum up a test taker's ability and proficiency in the areas of English language skills of non-native speakers; listening comprehension, structure and written expressions, reading comprehension, and writing skills (Chen, 2016). There are three kinds of TOEFL based on media it administers; iBT (Internet-based Test), CBT (Computer-based Test), and PBT (Paper-based Test). Nowadays, TOEFL iBT completely displaces CBT version and it turns out to be the most popular TOEFL test. In spite of the fact, many universities or colleges in Indonesia still keep on administrating the PBT version to test takers instead of CBT or iBT and Syiah Kuala University is one of them.

The language center of Syiah Kuala University holds the tests of TOEFL ITP and TOEFL Prediction, known as TOEFL-Like-Test and both are in forms of PBT. Many students within the university have enrolled in the test as a compulsory requirement to apply for final examinations; scrip, thesis, as well as dissertation. With various large group that take the tests every day it must be various results in TOEFL scores coming from the test takers. The newest regulation of the rector of Syiah Kuala University indents that all of the students from every faculty; except the English Study Program, must be taught about TOEFL skill in MKWU (Mata Kuliah Wajib Umum) Bahasa Inggris.

Furthermore, the capability of the students in TOEFL proficiency is surely not the same. This dissimilar might come from several factors. Zoghi et al. (2013) outline that language proficiency, courage, and learning styles are solely effect on test takers or learners ability in mastering a second language. In addition, gender has been spotted as an essential influencing factor in interfering second language acquisition (Zoghi et al. 2013). Males and females basically share different pattern in functioning the brain; females are more likely to use left brain, whilst males tends to specialize the right one (Banic, 1997). Some other researchers such as Halpern (1986), Collins et al. (2000), Swiatek & Lupkowski-Shoplik (2000) urge that gender differences positively impact learners' academic interest, need, and achievement.

### 3. Research Method

The design of this research is descriptive quantitative method. Given (2008) argues that quantitative research defines as a systematic and empirical procedure of investigation in which numerical data are magnified to gather information about an observable circumstance. The researcher, in this study, deployed some formulas as an attempt to investigate and compare the TOEFL scores between male and female students. To find the average score of the two groups the researcher magnified two mean scores ( $\bar{x}_1$ ) and ( $\bar{x}_2$ ), two variances ( $S_1$ ) and ( $S_2$ ), and *t-score*. The sample of this study were 10 males of undergraduate and 10 females of the same level, so there were 20 samples altogether from non-English majors of Syiah Kuala University. Those samples were selected by using simple random sampling.

The instruments used to gain the raw data or information were two sets of TOEFL PBT items. The first one was listening comprehension in which 25 questions administered to the samples. The second was structure and written expressions that consist of 25 questions. Each item scored 2 for the right answer, thus the highest score was 100. All of the items were adopted from the book of Longman Preparation Course for the TOEFL Test.

### 4. Results and Discussion

The results of the TOEFL PBT test of both male and female students are presented below:

**Tabel 1.**

No	Male Students	Female Students
1.	Number of sample ( $n$ ) = 20	Number of sample ( $n$ ) = 20
2.	The sum of males' scores ( $\sum X_1$ ) = 665	The sum of females' scores ( $\sum X_2$ ) = 575
3.	The sum of score squared of male students ( $\sum X_1^2$ ) = 53175	The sum of score squared of female students ( $\sum X_2^2$ ) = 36525

Those above total scores then analyzed by using some formulas such as Mean, Variance, homogeneity of variance, degree of freedom, and  $t_{\text{test}}$ . All of the data which have been measured are presented in this following summary table.

**Table 2.**

No	Statistical Symbols	Results of TOEFL PBT Test	
		Male Students	Female Students
1.	$\bar{x}$ (mean score)	33,25	28,75
2.	$S^2$ (variance)	1634,93	1052,20
3.	$F_{\text{obtained}}$ (homogeneity of variance)		1,55
4.	$F_{\text{table}}$		3,59
5.	$df_1$ & $df_2$ (degree of freedom)		2 & 17

6.	$\alpha$ (level of significance for two tailed-test)	0,05
7.	$t_{\text{test}}$	0,38
8.	$t_{\text{table}}$	1,725
Conclusion = $t_{\text{test}} < t_{\text{table}}$		

Based on the statistical results above, the mean score for the male students or  $\bar{x}_1$  is 33,25, whilst for the female students is  $\bar{x}_2$  is 28,75. It means that the mean score for the male students is higher than for the female students since the individual scores of male students are more than of female students. The highest score achieved by male students was 90, whereas the score of 80 was for the females. On the other hand, the lowest score for male and female student was the same; 30.

Next, the variance for both genders are dissimilar as well with 1634,93 and 1052,20 respectively. In other words, it shows how all of the scores spread out and describes the gap between the highest and the lowest marks. In addition, to determine whether the samples came from that population that has a homogenous variance, the researcher used the test of homogeneity of variance. The homogeneity of variance was analyzed by using these hypotheses: If  $F_{\text{obtained}} < F_{\text{table}}$  = the variance is homogeneous and If  $F_{\text{obtained}} > F_{\text{table}}$  = the variance is not homogeneous. After calculating the degree of freedom of both genders;  $df_1 = 2$  and  $df_2 = 17$ , it can be spotted in the  $F_{\text{table}}$  for the degree of freedom at the level of significance ( $\alpha$ ) = 0,05 is 3,59. Hence, it can be learned that  $F_{\text{obtained}}$  (1,55) is lower than  $F_{\text{table}}$  (3,59), so the variance is homogeneous.

After all of the statistical symbols deployed and calculated, the last formula to be tabulated is that of  $t_{\text{test}}$ . The  $t_{\text{test}}$  was magnified as an attempt to find out whether gender issue might make a significant difference in the results of TOEFL test proficiency. The criteria of the  $t_{\text{test}}$  at the level of significance ( $\alpha$ ) 0,05 are : If  $t_{\text{test}} < t_{\text{table}}$ , it means that there is no significant difference in the TOEFL proficiency between male and female students and if  $t_{\text{test}} > t_{\text{table}}$ , it covers that gender issue does not merely effect the students' proficiency on the TOEFL test. From the above table, it can be noted that  $t_{\text{test}} < t_{\text{table}}$  or 0,38 lower than 3,59. It means that there is no significant difference in TOEFL proficiency among male and female students.

This result is in line with the prior research conducted by Lin & Wu (2003). They conducted a study to analyze whether gender variable interfere the proficiency in TOEFL test and the result unveils that the gender does not influence the proficiency. The prior theories of Second Language Acquisition (SLA) proposed by some researchers such as Burstall (1975) and Ehrlich (2001) claims that female learners are marked as the superior one in absorbing the second language compared to the male learners. This assumption positively contradicts with this recent result which points out that both male and female students do not share different proficiency in TOEFL test.

In a linear fashion, Brimer (1969) and Boyle (1987) also perceive that gender does not play a significant role in interfering second language acquisition. In their research it shows that, although male students seem more consent in listening comprehension compared to the females, there is no significant final result in it. Hence, they end up with the conclusion that there is no gender impact on second language proficiency.

## 5. Conclusions

Gender issue has been debatable variable in second language acquisition and learning. Many researchers undergone the study under the topic of it still obtain different data. Some researchers perceive that female learners show a greater superior in absorbing and actualizing second language. Meanwhile, other contradictive researches reveal that ultimately, gender does not influence the second language proficiency including TOEFL test which has been widely and officially noted as a high-stake test of second language test around the globe. However, this recent study adds one more relevant study to whom they believe gender does not make significant difference in proficiency of the TOEFL test. This research result shows that there is no significant different in TOEFL proficiency between male and female students. Although the result of raw score of male students shows that they achieved higher score in TOEFL test it ends up with a conclusion, after tabulating all of the data in the  $t_{\text{test}}$  to seek its significant difference, it statistically reveals that gender actually does not interfere the male and female students in achieving their proficiency in TOEFL test after all.

Moreover, this study has several limitations. Firstly, in term of number of participants it was only 40 respondent have been randomly selected. Secondly, number of TOEFL questions was limited; there were only 50 questions altogether. Finally, a narrow scope of research location is being another limitation of this research. The researcher underwent this research only in Syiah Kuala University due to the reason of limited time. Hence, it is strongly suggested to the further researchers to expand and explore more to gain a more generalizable result.

## References

- [1] Baldiga, K. (2014). Gender differences in willingness to guess. *Manag Sc*, 60 (2), 434-448.
- [2] Banich, M. T. 1997. Breakdown of executive function and goal-directed behavior. In M.T Banich (Eds). *Neuropsychology: The neural bases on mental function* (pp. 369-390). Boston, MA: Houghton Mifflin Company.
- [3] Bicap, B. (2013). Scale for Test Preparation and Test Taking Strategies. *Educational Science: Theory and Practice*, 13 (1), 279-289.
- [4] Brown, Douglas, (2007). *Teaching by principles, An interactive approach to language pedagogy*. United State of America: Pearson Longman.
- [5] Burstall, C. 1975. Factors affecting foreign-language learning: A consideration of some relevant findings. *Language Teaching and and Linguistics Abstract*, 8 (5), 125.

- [6] Chen, A.C. 2016. A critical evaluation of text difficulty development in ELT textbook series: *A corpus-based approach using variability neighbor clustering*. *System*, 58, 64-81.
- [7] Cole, N.S. 1997. *The ETS gender study: How females and males perform in educational setting*. Princeton, NJ: ETS.
- [8] Collins, C., Kenway, J., & Mcleod, J. 2000. *Factors influencing the educational performance of males and females in school and their initial destinations after leaving school*. Canberra: commonwealth of Australia.
- [9] Dodeen, H.M., Abdelfatah, F., & Alshumnari, S. 2014. Test-Taking Skills of Secondary Students: The Relationship between Motivation, Attitudes, Anxiety, and Attitudes toward Tests. *South African Journal of Education*, 34 (2), 1-8.
- [10] Ellis, A. P. J., & Ryan, A. M. 2003. Race and Cognitive Ability Test Performance: The Mediating Effects of Test Preparation, Test-Taking Strategies Use and Self Deficiencies. *J Appl Soc Psychol*, 33 (12), 2607-2629.
- [11] Ehrlich, S. 2001. *Representing Rape: Language and Sexual Consent*. London: Routledge.
- [12] Given, Lisa M. 2008. *The SAGE Encyclopedia of Qualitative Research Methods*. Los Angeles: SAGE Publications. ISBN 1-4129-4163-6.
- [13] Halpern, J. Y. 1986. Reasoning about knowledge: An overview. In Halern, JY (Eds). *Proceeding of the 1986 Conference of theoretical Aspects of reasoning about Knowledge*, pp 1-18. San Mateo, CA: Morgan Kaufmann Publishers.
- [14] Hong, E., Sas, M., & Sas, J. T. 2006. Test Taking Strategies of High and Low Mathematics Achievers. *J Educ Res*, 99 (3), 144-155.
- [15] Hyde, J.S & Linn, M.C. 1998. Gender Differences in Verbal Ability: A meta-analysis. *Psychological Bulletin*, 104(1), 53-69.
- [16] Lee, Saeun. 2018. *Improvement of pre- and post-tests and gender differences on TOEFL Scores*. Retrieved from <https://miyazaki-mu.repo.nii.ac.jp>
- [17] Lin, J., & Wu, F. 2004. Differential performance by gender in foreign language testing. *Paper presented at the annual meeting of the national council on measurement in education*, Chicago, IL.
- [18] Mahmud, M. (2014). The EFL students' problems in answering the Test of English as a Foreign Language (TOEFL): A study in Indonesian context. *Theory and Practice in Language Studies*, 4(12): 2581-2587.
- [19] Munadia. 2016. *An analysis of EFL students' difficulty in the TOEFL structure and written expression section*. Thesis. Department of English language education, UIN Arraniry.
- [20] Naylor, F. D. (1997). Test-taking anxiety and expectancy of performance. In J. P. Keeves (Ed.), *Educational Research, methodology, and measurement: An international handbook* (2nd ed., pp. 971–976). Cambridge: Cambridge University Press.
- [21] Nurhayati, I.K., & Gir, R. R. W. 2014. Analisis perbandingan nilai TOEFL dengan nilai mata kuliah bahasa Inggris (Analysis of Comparison between

- TOEFL Score and the score for English module. *Jurnal Socioteknologi*, 13 (2), 134-146.
- [22] Phoham, James. 1999. *Using Standards and Assessment*. Los Angeles: Association for Supervision and Curriculum Department.
- [23] Stendlund, T., SundstÖm, A., & Jonsson, B. (2017). Effect of Repeated Testing on Short-hand Long Term Memoryperformance across Different Test Formats. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 36 (10), 1710-1727.
- [24] Stenlund, T., Sundström, A., & Jonsson, B. 2016. Effects of repeated testing on shortand long-term memory performance across different test formats. *Educational Psychology: An International Journal of Experimental Educational Psychology*, 36(10), 1710–1727. doi:10.1080/01443410.2014.953037.
- [25] Sternberg, R. J. 1998. *Metacognition, abilities, and developing expertise: What makes an expert student? Instr Sci*, 26, 127–140.
- [26] Sugeng, B., Saleh, S., & Suharto. 2012. Penguasaan Bahasa Inggris mahasiswa baru UNY tahun akademik 2005/2006-2009/2010 pada *criteria TOEFL-Like*. *LITERA*, 11 (2), 189-203.
- [27] Swiatek, M.A. & Lupkowski-Shoplik, A.E. 2000. Gender differences in academic attitudes among gifted elementary school students. *Journal for Education of the Gifted*. 23(4), 360-377.
- [28] Warfield, W., Larabee, R., & Geyer, R.W. (2013). Examining Results and Establishing Benchmark Data from the TOEFL ITP Test. *American Academic & Scholarly Research Journal*, 5, 3:191-198.
- [29] Zoghi, Masoud., Kazemi, Seyyed Ali., & Kalani, Ali. 2013. The effect of gender on language learning. *Journal of Novel Applied Sciences*, 2(S4), 1124-1128.