

[Article]

Predicting English Language Proficiency Scores among Study-Abroad Students in a University Setting

Ryan BARNES

Faculty of Economics
Nagoya Gakuin University

Abstract

Demographic factors, such as learner background variables, are strongly correlated with aptitude test scores of international English as a Second Language learners. This study conducted a multiple linear regression to assess the role of gender, length of study abroad, and age of international college students enrolled in an English for Academic Purposes program at a Midwestern university in the United States (n=626) in predicting scores on a test of English proficiency. Regression results were significant, although the explanatory power of the overall model was weak. These results suggest that certain learner backgrounds play a role in predicting test performance, and may be of interest language program designers and classroom teachers in placing students, and planning for English as a Second Language courses.

Keywords: ●●●, ●●●,



バーンズ ライアン

名古屋学院大学経済学部

発行日 2023年10月31日

Introduction

Although cognitive factors (Cross, 2010; Nakata, 2008; Zarei & Elekaei, 2013) and psychological factors (Engin, 2009; Masgoret & Gardner, 2003; Sampson, 2015) have been extensively studied in the literature for second language acquisition, recent studies (Olivares, 2012; van der Slik, van Hout, & Schepens, 2015; e.g., Vandergrift & Baker, 2015) have called for examining the role of learner background variables as well. Learner backgrounds are a demographic subset of learner variables and can include details such as gender, age, native language, and length of time abroad in an English-speaking environment.

The Michigan Test of English Language Proficiency (MTELP Series) is a test to measure language achievement and language acquisition of English language learners studying in an institutional language program. The exam is often used in determining placement for students in an appropriate English as a Second Language (ESL) level (Hille & Cho, 2020; Walter & Hentschel, 2014; Zhang, 2010), and tests listening comprehension, reading comprehension, grammatical knowledge and vocabulary range. This study will examine the role that learner background variables have in predicting the weighted, cumulative, converted score of this test as a holistic score of language proficiency.

Literature review

In language learning contexts, learner background variables refer to individual characteristics and experiences that can influence language learning. In the literature, these variables frequently overlap with individual differences (Gradman & Hanania, 1991; Oxford & Ehrman, 1992), but often center on non-changeable characteristics as age, native language, gender, educational history, and length of stay in a country where the target language is spoken (Spurling & Ilyin, 1985).

Learner background variables that have been found to be associated with proficiency outcomes in second language development include gender, where there was a gender gap between females and males in both written and spoken second language proficiency (Agustín Llach & Terrazas Gallego, 2012; Canga Alonso, 1970; Nyikos, 1990); study abroad experience, where students made significant gains in a second language environment compared to their peers who studied at their home university (Llanes & Muñoz, 2009; Segalowitz et al., 2004; Vande Berg, Connor-Linton, & Paige, 2009); and age (Nikolov & Djigunovic, 2006; Singleton, 2001), where there is a hotly debated topic that younger learners learn second languages better than older learners (e.g., Muñoz & Singleton, 2011). In this study, I hypothesize that the three variables of length of study abroad experience, gender, and age would significantly predict test scores, in line with the previous studies mentioned above.

Research Question

- RQ: Do gender, length of study abroad experience, or age predict test scores for a standardized test for English language proficiency?

Methodology

Data was cloned from the University of Pittsburgh English Language Institute Corpus (PELIC) GitHub repository, “a large learner corpus of written and spoken texts ...collected in an English for Academic Purposes (EAP) context over seven years in the University of Pittsburgh’s Intensive English Program, and were produced by students with a wide range of linguistic backgrounds and proficiency levels” (Juffs, Han, & Naismith, 2020). The combined dataset includes anonymized learner background variables and test scores for 625 students.

Measures

The MTELP Combined Score is the weighted, cumulative, converted score of the MTELP Series and was used as the outcome variable. The converted scale ranges from 0 to 100 points. The predictor values i.e., Gender, Years Abroad, Student Age were all self-reported. Gender included only binary male and female in this sample. Years Abroad was taken from a drop-down list indicating number of years in an English-speaking environment and included less than 1 year, 1-2 years, 3-5 years, and more than 5 years. Student Age at the time of enrollment was measured in years.

Procedures

Data analysis was conducted using R version 4.3.1 (2023-06-16) (R Core Team, 2023) in RStudio version 2023.6.1.524 (Posit team, 2023) on a computer running macOS version 13.1. Data cleaning was completed and missing data were removed using listwise deletion. Data importation, tidying, wrangling, and visualization were done using the tidyverse version 2.0.0 (Wickham et al., 2019) collection of R packages. Descriptive statistics and an assessment of assumptions were completed.

The following model equation was used:

$$\begin{aligned} MTELP = & \beta_0 + \beta_1 X_{Age} + \beta_2 X_{Gender} + \\ & \beta_3 X_{No\ Years\ Abroad} + \beta_4 X_{<1\ Year\ Abroad} + \beta_5 X_{1-2\ Years\ Abroad} + \\ & \beta_6 X_{3-5\ Years\ Abroad} + \beta_7 X_{>5\ Years\ Abroad} + \varepsilon \end{aligned}$$

Results

The data was analyzed for outliers and evaluated based on the assumptions of normality, linearity, homoscedasticity, independence, and multicollinearity, as well as specification and measurement error. While the assumptions of homoscedasticity, independence, linearity were deemed reasonably tenable, three outliers—identified by univariate and bivariate analysis—showed that two ages were unusually high (86 and 100 years old). It was suspected that the ages were misreported by the students—one birth year was entered as 19, possibly causing the data entry software to assign the maximum age of 100, and the other birth year was 1938, which may also be a data entry error, as it is not reasonable to expect an 86 year old to participate in a university-based study abroad program. As a result, these two cases were deleted due to suspected inaccuracies self-reporting. A third outlier was a MTELP Combined Score that was scored as 1. As this score is the weighted, cumulative, converted score of a series of exams, scoring a 1 seemed highly unlikely in this context. Upon reviewing the scores from the series, this Combined Score was inconsistent with points scored on other exams. As a result, this case was deleted due to suspected inaccurate data entry.

Descriptive Statistics

Descriptive statistics can be found in Table 1. MTELP Total Combined Scores ranged from 20 to 88 with a mean of 55.85 (SD=13.69). Age, measured in years, ranged from 18 to 64 with a mean of 26.88 (SD=6.30). Years Abroad heavily skewed toward students who had experience abroad of less than one year, i.e., 414 of 625 students. As for self-reported genders, males ($n_m=333$) slightly outnumbered females ($n_f=292$).

Table 1 Summary statistics for each of the observed variables

	vars	n	mean	sd	median	trimmed	mad	min	max	range	skew	kurtosis	se
Gender*	1	625.00	1.53	0.50	2.00	1.54	0.00	1.00	2.00	1.00	-0.13	-1.99	0.02
Years Abroad*	2	624.00	2.39	1.14	2.00	2.25	1.48	1.00	5.00	4.00	0.92	0.17	0.05
Age	3	625.00	26.88	6.30	26.00	26.25	5.93	18.00	64.00	46.00	1.24	2.81	0.25
MTELP Total Combined Score	4	625.00	55.85	13.69	55.00	55.24	16.31	20.00	88.00	68.00	0.29	-0.69	0.55

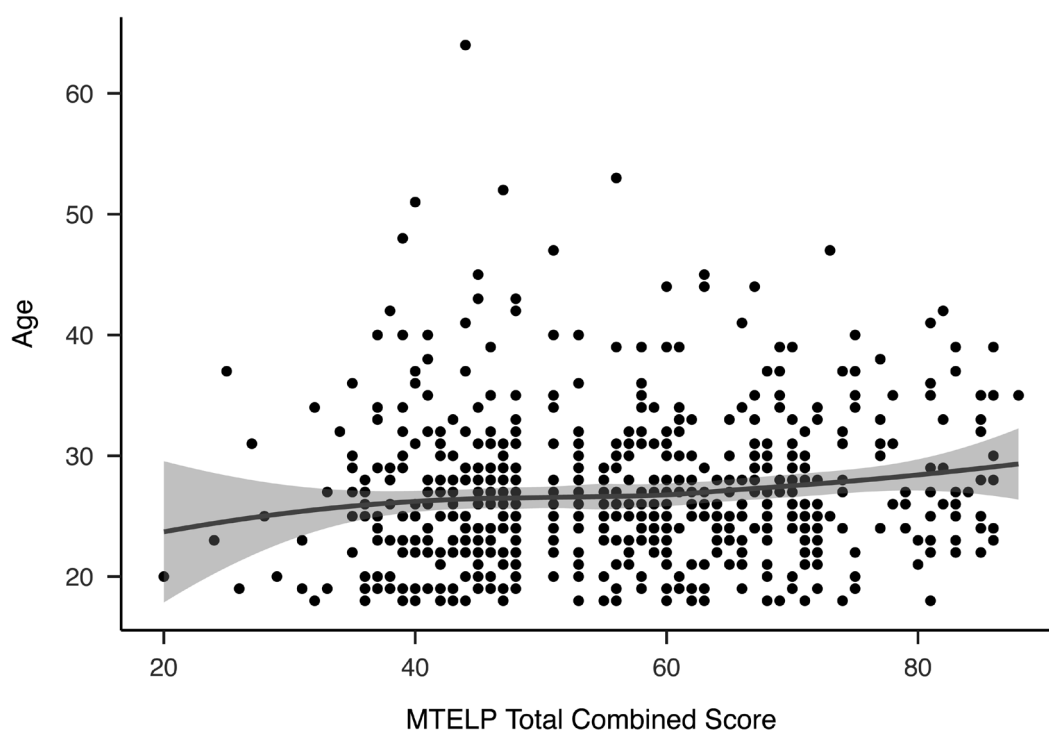


Figure 1 Scores by age

Model Results

Results of the multiple regression can be seen in Table 2. The overall model is statistically significant at $F(6, 617) = 5.32, p < .001$. The overall model fit indicated that the three predictor variables analyzed accounted for scores accounted for 4.92% of the variance in test scores ($R^2 = 0.05$). The results indicated that for each year of life, there was a 0.20 point increase in test scores ($p = 0.02$), controlling for all other predictors. Being female accounted for a 5 point increase in test scores ($p < .001$), controlling for all other predictors. None of the categories of years abroad significantly predicted a change in test scores, controlling for all other predictors.

Discussion

This study explored the relationship between several learner background variables and English language proficiency scores for students in a university study abroad program. While two variables were found to be significant predictors, the low explanatory power of the model indicates that there are factors or constructs influencing language proficiency. This specification error provides an opportunity for further research.

Table 2 Results of Multiple Regression Model

	<i>Dependent variable:</i>
	MTELP Total Combined Score
Age	0.202**(0.087)
Gender: Male	-5.005***(1.077)
Years Abroad: Less than 1 year	-0.139(1.442)
Years Abroad: 1-2 years	0.079(1.761)
Years Abroad: 3-5 years	0.636(2.378)
Years Abroad: More than 5 Years	2.665(2.183)
Constant	52.811***(2.668)
Observations	624
R ²	0.049
Adjusted R ²	0.040
Residual Std. Error	13.376 (df=617)
F Statistic	5.317***(df=6;617)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

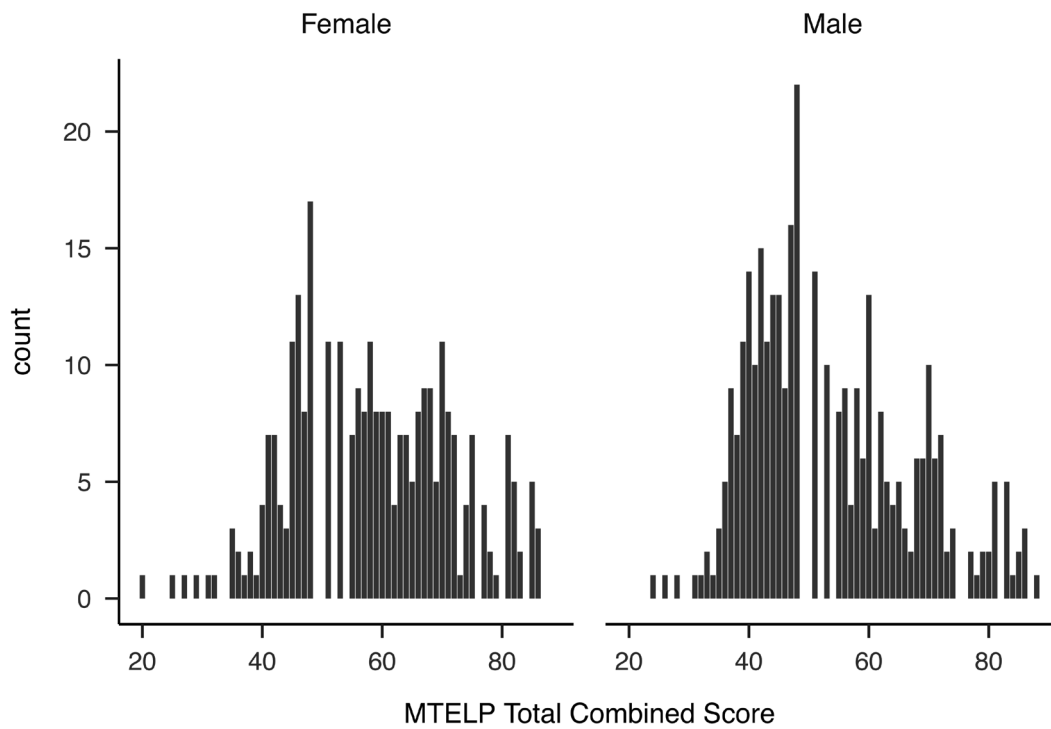


Figure 2 Frequency count for scores by gender

I hypothesized that females (see Figure 2) and older students (see Figure 1) would perform better on this language proficiency test, and the data in this sample supported that. There are a number of

debates in the literature that seek to explain the gender gap in second language acquisition, where it is hypothesized that an interaction of environmental factors such as learner strategies, motivation, orientation and nature-based abilities (e.g., van der Slik et al., 2015), or maturity levels (e.g., Agustín Llach & Terrazas Gallego, 2012). As for older students, reasons may include more experience studying the language, improved learner strategies, or the possibility of improving test scores based on having taken the test multiple times (Barkaoui, 2017).

Years abroad did not have much impact on the scores. This may be due to the clustering nature of self-reporting data in a range: a drop-down box, which separates experience into discrete categories, rather than having the participant input a number of years. A heavy majority of students (66%) selected “none” or “less than 1 year” as the length of their study abroad experience, possibly skewing the data. Although it is logical to expect that more time abroad in an English-speaking country would improve one’s English (Segalowitz et al., 2004), the scores remained consistent regardless of time abroad, as seen in Figure 3.

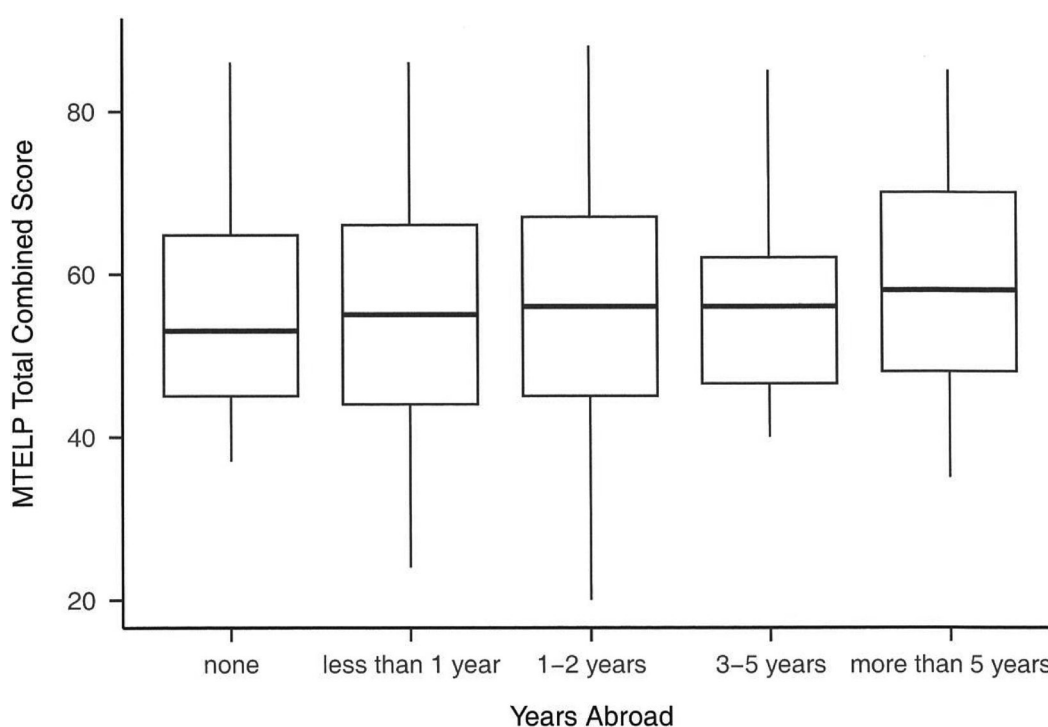


Figure 3 Scores by years abroad

This may also be due to the relatively young age of the participants, as few of them would have had the opportunity to study abroad for more than a few years by the time they reached the median age of 26 in this sample. Examining data from a wider variety of sources, for example ones that go beyond placement tests that are typically given to learners who are newer in an English-speaking

environment, may provide different results.

Conclusion

This study does not imply causation. Self-reporting of some data, and the limitations of entering that data (e.g., binary-only genders, clustering experience abroad into a range of years) may serve as a limitation of obtaining accurate results. To better understand the association between language proficiency test scores and learner variables, future researchers may wish to conduct qualitative research to gain a deeper understanding about the language learning experiences of participants during their time abroad. It may also be interesting to look at the learner variables of the students' native languages and countries, ways of study, or learning environments, for example. Finally, motivational factors are a very important element of language acquisition demographic variables (Engin, 2009; Masgoret & Gardner, 2003; Sampson, 2015), and these individual differences are essential in understanding how one learns a new language.

References

- Agustín Llach, M. P., & Terrazas Gallego, M. (2012). Vocabulary knowledge development and gender differences in a second language. *Estudios de Lingüística Inglesa Aplicada*, 12, 45–75.
- Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., ... Iannone, R. (2021). *Rmarkdown: Dynamic documents for R* [Manual]. Retrieved from <https://github.com/rstudio/rmarkdown>
- Barkaoui, K. (2017). Examining repeaters' performance on second language proficiency tests: A review and a call for research. *Language Assessment Quarterly*, 14(4), 420–431. <https://doi.org/gnzpc4>
- Canga Alonso, A. (1970). Receptive vocabulary of CLIL and non-CLIL primary and secondary school learners. *Complutense Journal of English Studies*, 23, 59–77. https://doi.org/10.5209/rev_CJES.2015.v23.51301
- Cross, J. (2010). Raising L2 listeners' metacognitive awareness: A sociocultural theory perspective. *Language Awareness*, 19(4), 281–297. <https://doi.org/10.1080/09658416.2010.519033>
- Engin, A. O. (2009). Second language learning success and motivation. *Social Behavior and Personality*, 37(8), 1035–1041.
- Gradman, H. L., & Hanania, E. (1991). Language learning background factors and ESL proficiency. *The Modern Language Journal*, 75(1), 39–51. <https://doi.org/10.1111/j.1540-4781.1991.tb01081.x>
- Hille, K., & Cho, Y. (2020). Placement testing: One test, two tests, three tests? How many tests are sufficient? *Language Testing*, 37(3), 453–471. <https://doi.org/10.1177/0265532220912412>
- Juffs, A., Han, N.-R., & Naismith, B. (2020). *The University of Pittsburgh English Language Institute Corpus*

Predicting English Language Proficiency Scores among Study-Abroad Students in a University Setting

- (*PELIC*). Zenodo. <https://doi.org/10.5281/ZENODO.3991977>
- Llanes, À., & Muñoz, C. (2009). A short stay abroad: Does it make a difference? *System*, *37*(3), 353–365. <https://doi.org/10.1016/j.system.2009.03.001>
- Masgoret, A. M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and associates. *Language Learning*, *53*(1), 123–163.
- Muñoz, C., & Singleton, D. (2011). A critical review of age-related research on L2 ultimate attainment. *Language Teaching*, *44*(01), 1–35.
- Nakata, T. (2008). English vocabulary learning with word lists, word cards and computers: Implications from cognitive psychology research for optimal spaced learning. *ReCALL*, *20*(01). <https://doi.org/10.1017/S0958344008000219>
- Nikolov, M., & Djigunovic, J. M. (2006). Recent research on age, second language acquisition, and early foreign language learning. *Annual Review of Applied Linguistics*, *26*, 234–260. <https://doi.org/10.1017/S0267190506000122>
- Nyikos, M. (1990). Sex-related differences in adult language learning: Socialization and memory factors. *The Modern Language Journal*, *74*(3), 273–287. <https://doi.org/10.1111/j.1540-4781.1990.tb01063.x>
- Olivares, G. (2012). Foreign language learner variables in an urban commuter university setting. *Arizona Working Papers in SLA & Teaching*, *19*, 1–23.
- Oxford, R. L., & Ehrman, M. (1992). Second language research on individual differences. *Annual Review of Applied Linguistics*, *13*, 188–205. <https://doi.org/10.1017/S0267190500002464>
- Posit team. (2023). *RStudio: Integrated Development Environment for R*. RStudio, PBC. Retrieved from <http://www.posit.co/>
- R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing.
- Sampson, R. J. (2015). Tracing motivational emergence in a classroom language learning project. *System*, *50*, 10–20. <https://doi.org/10.1016/j.system.2015.03.001>
- Segalowitz, N., Freed, B., Collentine, J., Lafford, B., Lazar, N., & Díaz-Campos, M. (2004). A comparison of Spanish second language acquisition in two different learning contexts: Study abroad and the domestic classroom. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *10*(1), 1–18. <https://doi.org/ghfp6z>
- Singleton, D. (2001). Age and second language acquisition. *Annual Review of Applied Linguistics*, *21*, 77–89. <https://doi.org/b6sfxv>
- Spurling, S., & Ilyin, D. (1985). The impact of learner variables on language test performance. *TESOL Quarterly*, *19*(2), 283. <https://doi.org/10.2307/3586830>
- van der Slik, F. W. P., van Hout, R. W. N. M., & Schepens, J. J. (2015). The gender gap in second language acquisition: Gender differences in the acquisition of Dutch among immigrants from 88 countries with 49 mother tongues. *PLOS ONE*, *10*(11), e0142056. <https://doi.org/f799n5>
- Vande Berg, M., Connor-Linton, J., & Paige, R. M. (2009). The Georgetown Consortium Project: Interventions for Student Learning Abroad. *Frontiers: The Interdisciplinary Journal of Study Abroad*, *18*(1), 1–75. <https://doi.org/10.36366/frontiers.v18i1.251>
- Vandergrift, L., & Baker, S. (2015). Learner variables in second language listening comprehension: An exploratory path analysis. *Language Learning*, *65*(2), 390–416. <https://doi.org/f7cbtk>

- Walter, D., & Hentschel, J. (2014). Dimensionality and factor structure of an English placement test. *Cambridge English: Research Notes*, (55), 34-45. Retrieved from https://michiganassessment.org/wp-content/uploads/2020/02/20.02.pdf.Res_.DimensionalityandFactorStructureofanEnglishPlacementTest.pdf
- Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L. D., François, R., ... Yutani, H. (2019). Welcome to the tidyverse. *Journal of Open Source Software*, 4(43), 1686. <https://doi.org/ggddkj>
- Zarei, A. A., & Elekaei, A. (2013). The effect of attitude on L2 learners' choice of compensation and meta-cognitive strategies. *International Journal of Applied Linguistics & English Literature*, 2(1), 186-193.
- Zhang, B. (2010). Assessing the accuracy and consistency of language proficiency classification under competing measurement models. *Language Testing*, 27(1), 119-140. <https://doi.org/10.1177/0265532209347363>