

Exploring Learner Proficiency Test Scores through Cluster Analysis

Kimberly Klassen and Tetsuya Fukuda
English for Liberal Arts
International Christian University

Abstract

This paper reports on the latest development of an ongoing study investigating differences in L2 English proficiency after one year in an intensive English program at a small, private university in Japan. Students' English proficiency were measured using the IELTS test at the end of the academic year. Volunteer participants answered a survey about their preparation and experience with the exam. Prior investigation through interviews with participants who improved their scores and those who were less successful revealed much variation in terms of test preparation and prior test experience. Thus, for the current study, cluster analysis was used to explore and identify subgroups based on their test preparation and prior test experience. After the test scores and the responses to Likert-scale question items were standardized with *z*-scores, subgroups among the larger cohort ($N = 183$) were identified using cluster analysis. The analysis revealed six subgroups, each distinctly characterised by their IELTS test scores, test familiarity and preparation, and perceived future need for English. Because the cluster analysis was successful in revealing unforeseen patterns, future directions will include using the procedure to explore proficiency gains with factors of test preparation, test experience and perceived future use of English.

Why some learners make large proficiency gains while others make little or no gains in the same period is of great interest to classroom teachers, program administrators, and SLA researchers alike. Having a better understanding of the factors, or combination of factors, that contribute to learners' success could enable various stakeholders to maximize conditions for improving learning outcomes. The study described here is the latest installment in ongoing research conducted in an English program for undergraduates at a private Japanese university. The research began in 2012, building on findings from previous studies. Prior findings indicated that the number of classes students took in the program was a predictor of their proficiency gains, while study abroad and prior academic writing experience were not (Erdelyi et al., 2018). Survey responses and follow-up interview data revealed much variation in terms of motivation, test preparation and study abroad experience among the learners that were successful and those that were less successful (Yagi & Fukuda, 2020). Due to this large variation among participants, cluster analysis was used in the present study to explore new ways of grouping the learners, and thus potentially learn more about combination of factors that predict success.

Cluster analysis is a "multivariate exploratory procedure that is used to group cases" and "is particularly relevant when there is evidence to suggest that different subgroups of learners may utilize different pathways to language learning, including different strategies, aptitudes, [and] motivational profiles" (Staples & Biber, 2015, pp. 243, 244). While cluster analysis is not common in second language research (Csizér & Dörnyei, 2005), use of the procedure has been increasing, in particular to investigate the role of individual differences on language learning (Staples & Biber, 2015). As Skehan (1986) explains in an early study employing cluster analysis, while there is value in more commonly used statistical techniques,

useful information pertaining to the individual learner may be missed. For example, regression analysis assumes a linear relationship between two variables; however, such an analysis would conceal important information relating to the threshold levels of certain abilities necessary for success. In this regard, cluster analysis has the potential to offer more information about individual differences than traditional techniques (Skehan, 1986, p. 82). Additionally, Alexander and Murphy (1999) summarize three advantages of cluster analysis over other traditional statistical analyses employed for categorizing techniques. First, unlike other multivariate techniques, such as confirmatory factor analysis, cluster analysis does not require the data to be normally distributed. Second, calculations employed to determine the clusters are relatively simple and acceptable from a mathematical perspective. Finally, the order of variables put in calculation does not affect the outcomes.

An illustration of how cluster analysis can be used as a complementary statistical procedure is found in Gu and Johnson's (1996) investigation of vocabulary learning strategies and learning outcomes. The authors collected data with a questionnaire to gauge participants' ($n = 486$) beliefs and approaches toward vocabulary learning, along with general English proficiency and vocabulary size scores. The authors performed correlation analysis between the strategies used, and vocabulary size and general proficiency levels; they also did multiple regression to determine the best predictors. Finally, they used cluster analysis to classify participants by their strategy use and learning outcomes. Gu and Johnson (1996) identified five groups in their analysis, each group distinct in terms of their strategy use and proficiency gains. The authors were able to conclude that differences in proficiency gains and vocabulary size were a function of strategy use and beliefs, and not participants' initial proficiency levels, as there was no statistically significant variation between the clusters (Gu & Johnson, 1996, p. 666).

Some studies have used cluster analysis to explore L2 learner motivation in Japan. For example, Tsuda and Nakata (2013) investigated motivation as an aspect of self-regulation among Japanese high school students ($N = 1076$). The authors first used exploratory factor analysis and identified five significant factors. Then, they conducted a cluster analysis to group students into four distinct profiles. Finally, they conducted interviews with representative learners from each group, in order to look more closely at learner background and experience. Tsuda and Nakata (2013) concluded that the cluster analysis results enabled them to identify complex factors influencing language learning, and that categorising learners as simply successful or unsuccessful is too simplistic (p. 85). In another study, Hiromori (2009) used structural equation modeling and cluster analysis to investigate the phases of learner motivation, specifically, the relationship between learners' intentions and their action control strategies. He categorized the university participants ($N = 148$), identifying four groups with distinct characteristics for each phase (pre- and post-) in the motivation model. The cluster analysis revealed important patterns: those groups that scored relatively high on intentional variables also scored higher on their actual engagement in the learning activities; those groups that scored middling or lower scores on the intentional variables also scored lower on engagement variables. These findings suggest that learners' intention had an influence on their actual learning behaviour (Hiromori, 2009, p. 319).

Given the potential for unforeseen information that cluster analysis can offer regarding individual differences in language learning, it was decided to use the statistical procedure as the next step in this longitudinal study. As much variation existed among the participants with regard to prior test experience and test preparation, and their proficiency scores, the focus for this initial analysis was on those predictors. Our research question is: What patterns can be identified among participants in terms of proficiency test scores, test experience, test preparation and extrinsic motivation?

Method

Participants

Participants ($N = 183$) were first-year university students who completed a one-year intensive English program at a university in Tokyo, Japan. A total of 363 students took the IELTS exam in March 2021 and were asked to participate in the survey; 51.8% of them responded. The data from five students with missing data were eliminated, and as a result, data of 183 students were kept. Their overall IELTS scores ranged from 4.5 to 8.0. According to IELTS (2020), those with 4.5 are described as “limited users” and those with 8.0 are “very good users.”

Data Collection and Analysis

The data analyzed in this study were twofold: IELTS overall scores and responses to the survey. Table 1 shows the five items and the descriptions of the options in the survey. Items 1–3 relate to the extent to which they prepared for the test (Prepare), Item 4 inquires about their prior experience of taking IELTS (Experience), and Item 5 inquires about how likely they are to use the score to study abroad (Future Use).

Table 1
The Question Items and the Options in the Survey

Question Item	Option 1	Option 2	Option 3	Option 4	Option 5
1 Did you prepare before you took the IELTS exam on March 13?	Very much	Quite a lot	Some	Not much	No
2 How many hours did you study to prepare for the March 13 IELTS exam?	More than 10 hours	Between 5 and 9 hours	Between 2 and 5 hours	Between 1 and 2 hours	None
3 Did you participate in the information sessions for IELTS held at ICU (Guidance Seminar on December 18 and Preparation Seminar on January 23)?	Both days	January only	December only	Part of one day	No
4 Did you take IELTS before you took it on March 13?	4 or more times	3 times	2 times	1 time	Never
5 How likely are you to use the IELTS results to study abroad in the near future, such as on an exchange program or graduate school?	Very likely	Quite likely	Somewhat	Not very likely	Not at all

The four factors analyzed in this study were IELTS scores, test preparation, test experience, and future use. To profile participants who exhibit the characteristics of similar combinations of these factors, cluster analysis was employed (Ward’s method with squared Euclidean distance technique), using SPSS 27 (IBM, 2020). The participants’ scores of IELTS and the survey results (Prepare, Experience, and Future Use) were used as clustering measures. Then, ANOVAs were performed to confirm the validity of the grouping and any significant differences were subjected to a post-hoc test. In calculating effect sizes in eta-squared, the criteria proposed by Plonsky and Oswald (2014) were employed, with $\eta^2 = .2$ representing a small effect, .5 medium, and .8 a large effect.

Results

To answer the research question (i.e. What patterns can be identified among participants in terms of proficiency test scores, test experience, test preparation and extrinsic motivation?), a cluster analysis using the four indicators as clustering measures was conducted. The four clustering variables were IELTS scores (IELTS), preparation for the exam (Prepare), prior experience of taking IELTS (Experience), and the perceptions of future use of the score (Future Use). Since the indicators had different units, standardized values were used. In determining the number of clusters, two features in the output, the dendrogram plot and the fusion coefficient, were examined, following Staples and Biber (2015). Both the dendrogram and the fusion coefficients obtained from the analysis supported the solution of categorizing the 183 participants into six groups, which are represented in Figure 1. To confirm the validity of the grouping solution, ANOVAs were conducted. Results showed that significant overall differences among the clusters were confirmed for all the four indicators with $p < .001$.

Figure 1
Cluster Profiles on the Four Factors of Examination

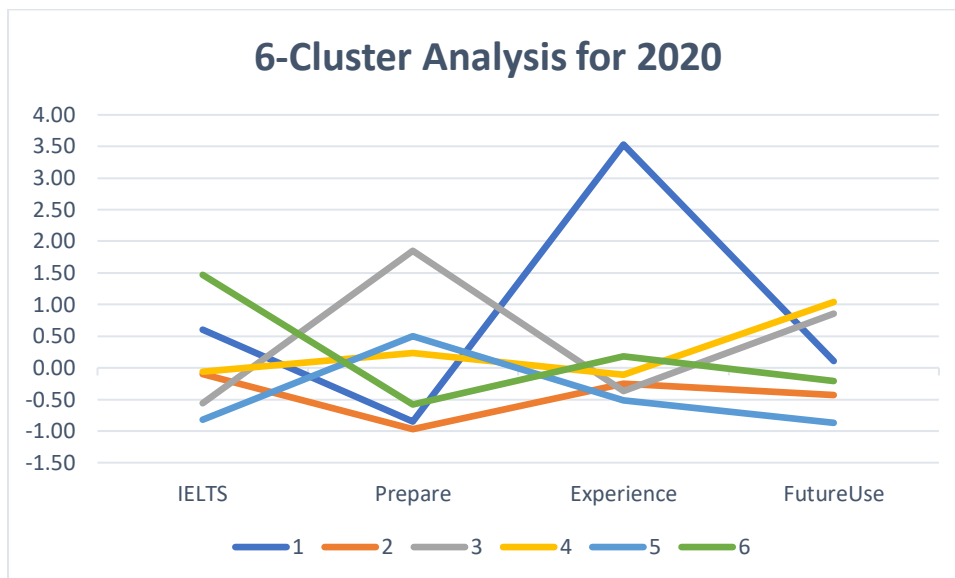


Table 2 displays the descriptive statistics for each cluster of participants with the results of ANOVAs and their effect size indicated in the eta-squared. All the ANOVA results were statistically significant ($p < .001$), indicating that the all groups were distinct from one another. The effect sizes were all medium according to the criteria in Plonsky and Oswald (2014). The findings revealed that the four variables were not equivalent among the six clusters of participants.

Exploring Proficiency through Cluster Analysis

Table 2
Means, Standard Deviations, and One-Way Analyses of Variance in the Four Factors

Cluster	1	2	3	4	5	6	<i>F</i> (5,182)	η^2
	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>	<i>M</i>		
	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)	(<i>SD</i>)		
IELTS	6.75 (.19)	6.22 (.07)	5.88 (.12)	6.25 (.09)	5.68 (.08)	7.42 (.08)	45.24***	.56
Prepare	1.93 (.22)	1.33 (.06)	4.17 (.10)	2.54 (.09)	2.80 (.11)	1.72 (.12)	64.62***	.65
Experience	5.00 (.00)	1.24 (.07)	1.13 (.09)	1.39 (.10)	1.02 (.02)	1.67 (.12)	118.34***	.77
Future Use	3.40 (.37)	2.78 (.14)	4.25 (.19)	4.45 (.08)	2.29 (.12)	3.03 (.12)	43.70***	.55

*** $p < .001$.

The following is a detailed description of the characteristics of each cluster. Cluster 1 consists of 10 students with the most experience of IELTS. All the students in this group had taken IELTS at least four times before. They did not prepare for the test probably because they were already familiar with the format. Their scores were relatively high and they are moderately interested in using the test score to study abroad. Cluster 2 represents the 41 students who did not prepare for the test, but their score was average. These students were not very interested in using the test score to study abroad. In Cluster 3 are the 16 students who prepared for the test a lot, taking preparation courses. The low experience level may have been behind the high rate of preparation. Their scores were not high, but they were interested in going abroad using the scores. Cluster 4 is a group of 44 students who stood out in their high interest in using the test score to go to a foreign university. The other factors, including test score, test preparation, and prior test experience, were all at a moderate level. Cluster 5 includes 42 students who were the least interested in using the score for studying abroad. Although they prepared for the test more than other students, their scores were lowest possibly because they had never taken the test before. Cluster 6 represents 30 students in this group that scored the highest scores in the test, but they did not prepare for the test much, their prior test experiences were just above the average, and they were not so much interested in using the score in the future.

Discussion

The cluster analysis revealed important information about the cohort that would not have been readily apparent through other statistical analyses. One interesting finding is that over a third (38.8%) of the learners taking the IELTS did not prepare much, had little experience with the test, nor did they identify a strong future need for the test score (represented by Clusters 2 and 6). The two groups differed in test scores: Cluster 2 had average scores while Cluster 6 had the highest scores among the cohort. Follow-up research is needed to understand why a significant portion of the cohort who did little preparation, had little test familiarity, and little future need for the score, decided to take it. Even though the learners did not have to pay for the exam, considering that taking the test represents a significant amount of time, it would be helpful to know more about their motivations.

Exploring Proficiency through Cluster Analysis

An important finding in terms of motivation was seen in Cluster 4 ($n = 44$), representing nearly a quarter of the cohort (24.0%). This group reported the strongest future need for an IELTS test score, and yet their preparation and experience with the test format was middling. This finding points to an opportunity within the program to better communicate with students who are highly motivated about the different ways they can prepare for and learn about the test format. Certainly, having experience with the test can be important, as indicated by Cluster 1. They reported the greatest familiarity with the test and also had above average test scores, though it is important to note this was a small group ($n = 10$).

Future research, of course, will also need to compare learning gains in order to account for differences in proficiency levels in the data. Prior research (Erdelyi et al., 2018; Yagi & Fukuda, 2020) had utilized program entrance and exit test scores in order to investigate reasons for learning gains (or losses). However, due to the Coronavirus pandemic, program entrance scores were not obtained for this cohort. For that reason, this cluster analysis focused only on the single IELTS test score. Conducting a future cluster analysis on more factors including scores from entrance and exit tests may reveal other types of learner clusters. In addition, analyses on learning gains may uncover more about the effects of test familiarity, test preparation and extrinsic motivations for test scores, and could potentially allow for stronger conclusions to be drawn regarding those effects. For example, in the current study, those learners who prepared the most for the test (Clusters 3 and 5) received the lowest test scores among the cohort. However, as little is known about their learning gains, it is impossible to draw any conclusion about the effect of test preparation. Clusters 3 and 5 were also the least experienced with the test, and this may have accounted for their higher preparation levels.

Conclusion

Much variation among cohorts was seen in prior research in terms of test preparation, test familiarity and motivation (Erdelyi et al., 2018; Yagi & Fukuda, 2020). For that reason, the current study employed cluster analysis in an attempt to learn about what factors, or combination of factors, might predict learning success. Data collection restraints (due to the Coronavirus pandemic) meant that the cluster analysis could only be performed using one test score (i.e., the program exit test score). Despite this restriction, the analysis still revealed information about the cohort that would not have been easily discoverable through other data collection means. The cluster analysis revealed six distinct groups of learners based on the four factors of test scores, test preparation, test familiarity and future use for the test score.

Because the analysis was successful in terms of what was learned about the cohort, the next step in the research will be to apply cluster analysis on a future cohort's learning gains. When Coronavirus restrictions are lifted, program entrance and exit test scores will be collected. It will also be possible to add other factors to the analysis, in addition to the factors of preparation, experience and motivation. Then, learning gains (or losses) can be examined in light of these factors. Teachers and program administrators can draw on the findings to help students reach their English academic learning goals during the first-year program and beyond.

References

- Alexander, P. A., & Murphy, P. K. (1999). Learner profiles: Valuing individual differences within classroom communities. In P. L. Ackerman, P. C. Kyllonen & R. D. Roberts (Eds.), *Learning and individual differences* (pp. 413–431). American Psychological Association.
- Csizér, K., & Dörnyei, Z. (2005). Language learners' motivational profiles and their motivated learning behavior. *Language Learning*, 55(4), 613–659.
- Erdelyi, T., Fukuda, T., & Yagi, K. (2018). Analysis of study abroad summer program. *OnCUE Journal*, 11(2), 116–122.
- Gu, Y., & Johnson, R. K. (1996). Vocabulary learning strategies and language learning outcomes. *Language Learning*, 46(4), 643–679.
- Hiromori, T. (2009). A process model of L2 learners' motivation: From the perspectives of general tendency and individual differences. *System*, 37, 313–321.
- IBM. (2020). IBM SPSS statistics for Macintosh, Version 27.0 [computer software]. Armonk, NY: IBM Corp.
- IELTS. (2020). *IELTS: About IELTS*. <https://www.ielts.org/about-ielts/what-is-ielts>
- Plonsky, L., & Oswald, F. L. (2014). How big is “big”? Interpreting effect sizes in L2 research. *Language Learning*, 64(4), 878–912. <https://doi.org/10.1111/lang.12079>
- Skehan, P. (1986). Cluster analysis and the identification of learner types. In V. Cook (Ed.), *Experimental approaches to second language learning* (pp. 81-94). Pergamon Books.
- Staples, S., & Biber, D. (2015). Cluster Analysis. In L. Plonsky (Ed.), *Advancing quantitative methods in second language research* (pp. 243–274). Routledge.
- Tsuda, A., & Nakata, Y. (2013). Exploring self-regulation in language learning: a study of Japanese high school EFL students. *Innovation in Language Learning and Teaching*, 7(1), 72–88.
- Yagi, K., & Fukuda, T. (2020). Exploratory investigation into successful and less successful learners of English with interviews. *Language Research Bulletin*, 35, 84–94.