AN INVESTIGATION INTO INDONESIAN EFL UNIVERSITY STUDENTS’ SPEAKING ANXIETY

Mukhlash Abrar
Universitas Jambi, Indonesia
mukhlash.abrar@unja.ac.id

Abstract: The general purpose of the study was to investigate EFL learners’ speaking anxiety and its associated factors. More specifically, the study was purposed to examine the constructs of a speaking anxiety and to explore the model for predicting speaking anxiety based on gender, proficiency and class type. Therefore, explanatory factor analysis and multiple linear regression analysis were used. The result indicated that there are two factors in learners’ speaking anxiety, but they are difficult to name as a result of overlapping variables in each component. With regards to a multiple linear regression test, the finding shows that proficiency variable is the most significant factor for predicting the variation in speaking anxiety.

Keywords: speaking anxiety, factor analysis, multiple linear regression

INTRODUCTION
Successful foreign language learning requires that students accommodate both cognitive and affective domains in their learning. Anxiety, an affective factor in language learning, is a frequently observed issue which arguably influences students’ performance and their academic achievement (Chastain, 1975; Dordinejad and Ahmadabad, 2014; Gardner et al, 1997; Horwitz, 2001; Luo, 2013; Saito and Samimy, 1996; Scott, 1986) and can negatively impact the learning process (Arnold and Brown, 1999; Gardner and Maclntyre, 1993; Oxford, 1996). This phenomenon is often identified in speaking modules of English as a Foreign Language (EFL) (Horwitz et al, 1986; Palacios, 1998; Price, 1991) because students are required to think and give a response to the initiation almost simultaneously. Thus, this
quantitative study investigates and reports EFL learners’ speaking anxiety and its associated factors and aims to provide some practical suggestions for language tutors and meaningful recommendations for future research.

LITERATURE REVIEW
Foreign Language Anxiety
Anxiety is one of the most highly researched phenomena in psychology and education. It is defined, in simple terms, as the feeling of unease or fear and is often identified in many social and learning contexts. According to Hilgard, Atkinson, and Atkinson (1971), anxiety is regarded as a psychological construct, commonly described as a state of apprehension or worry, which is indirectly associated with an object. Furthermore, according to Chastain (1988), anxiety is a state of uneasiness and apprehension initiated by the anticipation of something threatening. Along similar lines, Rachman (2004:3) asserts anxiety as “tense, unsettling anticipation of a threatening but vague event; a feeling of uneasy suspense”. Within the foreign language context specifically, Horwitz, Horwitz, and Cope (1986: 128) define anxiety as “a distinct complex of self-perceptions, beliefs, feelings, and behaviours related to classroom language learning arising from the uniqueness of the (foreign) language learning process”. Horwitz (2001) further proves that foreign language learning anxiety has a negative impact on learners’ performance and academic achievement. Moreover, MacIntyre (1998) asserts that anxiety is a feeling, a worry and an emotional reaction which arises while learning or using a second language, and which negatively impacts learning process. These imply that anxiety is a psychological construct which causes individuals worry or fear of something which negatively influences either the process or the achievement of learning a foreign language.

Anxiety, broadly speaking, can be classified into three types: trait anxiety, state anxiety and situation-specific anxiety. However, Horwitz, Horwitz, and Cope (1986) and William and Andrade (2008)
argue that not all anxiety types apply specifically to the foreign language learning context. Ellis (1994) states that trait anxiety refers to the stable tendency to be nervous in many circumstances. In other words, it is a part of a person’s characteristics and hence is a difficult trait to shed. An individual whose trait is anxious tends to feel anxious in a number of situations. State anxiety, as the name implies, on the other hand, relates to an unstable feeling of anxiousness that arises in specific circumstances as a response to an external stimulus (MacIntyre and Gardner, 1989; Spielberger, 1983).

A person experiencing state anxiety will feel tension or worry because they are exposed to a particular situation which causes them stress, but it will fade when the threat disappears. Finally, situation specific anxiety, according to McIntyre and Gardner (1991), can be seen as trait anxiety limited to a given context. It is stable over time, but inconsistent across varying circumstances. That is, it is prompted only by a particular setting or situation, such as taking a test, public speaking or speaking in a foreign language. Given the features of situation-specific anxiety, MacIntyre and Gardner (1991) suggest that foreign language anxiety should be studied with situation specific measures. Horwitz et al. (1986) further support the theory that foreign language anxiety is classified as situation-specific anxiety.

As previously indicated, anxiety, in principle, is a response triggered by external threat. Therefore, it can be attributed to several factors in the context of language learning in general. Young (1991) argues the possibility that language anxiety may emerge from three aspects of learning: the teacher, the learner and the instructional practice. More specifically, he claims six interrelated factors as the causes of learning anxiety, namely: (1) personal and interpersonal anxiety; (2) learner beliefs about language learning; (3) instructor beliefs about language teaching; (4) instructor-learner interactions; (5) classroom procedures; (6) language tests.

In the foreign language context, Aida (1994), Casado and Dereshiwsky (2004), Horwitz et al (1986), MacIntyre and Gardner (1989), and Pappamihiel (2002) conclude that anxiety is the result of
three possible affecting factors: communicative apprehension (e.g. difficulty in understanding the teachers’ instruction and/or peers’ talk); fear of negative evaluation (e.g. fear of being corrected by teachers/peers); test anxiety (e.g. fear of failing the module/lesson). These three affecting factors are well-known sources of anxiety and are relevant to the discourse of foreign language anxiety. In the present study, two of the three constructs - communicative apprehension and fear of negative evaluation - have been explored in more depth, by using them as subscales within the questionnaire.

The overall discussion on the concept of foreign language anxiety suggests that anxiety is one of the most examined affective factors which may negatively influence foreign language learners’ performance and/or acquisition of the language due to their difficulty in understanding the lesson/instruction, worry of getting negative feedback from onlookers, or fear of failing the class. Since this psychological construct is personal, every foreign language learner may experience a different level of anxiety. Students with higher levels of anxiety had a weaker performance or achieved less compared to less anxious students (Hewitt and Stephenson, 2012; Kleinmann, 1977; Steinberg and Horwitz, 1986).

**Speaking and Foreign Language Anxiety**

Speaking is generally recognised as a fundamental language skill for effective interaction in any language, including for non-native language speakers. Its nature is exceptionally distinct from other language skills in that it is a verbal productive skill that requires the speakers’ mastery in linguistic and sociolinguistic competence (Nunan, 2003, 2009; Hinkel, 2005). Of the two competences, much of the research reported that linguistic competence appeared to be the one of the main challenges for non-native speakers in learning and/or practising a foreign language (Abrar and Mukminin, 2016; Al-Hosni, 2014; Al-Jamal & Al-Jamal, 2014; Arju, 2011; Gan, 2012; Keong et al, 2015; Lee, 2009; Priyatno, 2013; Paakki, 2013; Wang and Roopchund,
owing to the numerous components in which one must gain proficiency.

Syakur (1987) details that there are, at least, five components with which speakers need to engage when speaking, including: comprehension (ability to understand the meaning and the capacity to engage in a conversation); grammar (well-structured sentences); pronunciation (comprehensible articulation); word choice (appropriate diction); fluency (the ease and speed of the flow of expression). Syakur (1987) theory was used as the framework in constructing the questionnaire for this study.

In the context of foreign language learning, anxiety is often associated with speaking ability (Lucas, 1984; Phillips, 1992, Price, 1991). Many studies have consistently shown that foreign language anxiety and speaking ability are, to a certain extent, interwoven. Horwitz et al (1986), for instance, argued that communication comprehension is conceptually relevant to foreign language anxiety. Price (1991), then, revealed that speaking in front of peers in the target language provoked the most anxiety for learners in her foreign language class. On a similar thread, Palacios (1998) asserted that speaking caused the most anxiety among foreign language learners. Clearly, the correlation between speaking and anxiety is a cause-effect relationship in which speaking itself leads to anxiety for foreign language users/learners.

**Studies on Foreign Language Anxiety**

There has recently been a marked increase in studies of foreign language anxiety, specifically relating to speaking skills. Much of the research examined anxiety with its associated affecting factors/variables, such as gender and proficiency. Regarding gender proficiency, Çağatay (2015) researched four associated factors in speaking anxiety, of which gender was one. She administered an 18-item questionnaire to 147 students of an English preparatory programme of a state university in Turkey. Her findings suggested that there is a statistically significant difference between male and
female students’ anxiety in speaking a foreign language (English), noting that female students tend to be more anxious than male students.

Many other studies (Park and French, 2013; Ekström, 2013; Öztürk, and Gürbüz, 2013; Hsu, 2012; Tianjian, 2010; Occhipinti, 2009; Huang, 2004) supported this conclusion. A study conducted by Fariadian, Azizifar and Gowhary (2004), however, showed the opposite. They investigated gender contribution in EFL speaking anxiety among Iranian learners. With a total number of 80 participants involved in their study, they found that there is a significant difference between the two gender categories. Male respondents were reported to experience slightly higher levels of anxiety than females. Nevertheless, other research (Debreli and Demirkan, 2015; Şimsek, 2015; Muhaesin and Al-Haq, 2012; Cui, 2011; Aida, 2004; Voorhees, 1994) reported that there is no significant difference between male and female respondents in foreign language speaking anxiety.

With regard to the proficiency variable, the result of most related studies showed that more language-proficient participants tended to have a lower anxiety level when learning and speaking English than less-proficient participants, although the difference was not always significant (Çağatay, 2015; Tercan and Dikilitas, 2015; Zhao and Whitchurch, 2011; Tianjian, 2010; Liu, 2006). Tanjian (2010), for example, investigated Chinese EFL learners’ speaking anxiety including gender and proficiency differences. The researcher involved 240 participants and divided them into three categories of proficiency level.

The result of the research suggests that students from the lower proficiency group experienced more anxiety than the higher proficiency group. A study conducted by Debreli and Demirkan (2015), in contrast, produced a different result. By comparing anxiety between elementary and pre-intermediate level of EFL students, their findings suggest that more proficient EFL learners are more anxious in learning language than less proficient learners. They argue that as
the proficiency increases, the learners need to complete more demanding tasks and have more concerns about their teachers’ expectation.

In this present study, the researcher also used gender and proficiency as independent variables (IVs). Male and female respondents were used in order to satisfy the gender variable, while proficiency was subjectively measured and classified by looking at the respondents’ experiences in attending English courses. In addition, the researcher added another IV named “class type” which was divided into regular and regular-\textit{mandiri} categories. This variable is unique to the Indonesian university context whereby regular class students are admitted on the basis of passing a national entrance exam while regular-\textit{mandiri} class students are admitted on the basis of passing a university entrance exam.

The present study attempts to investigate Indonesian EFL learners’ speaking anxiety and its associated factors. The study addressed the following alternative hypotheses to be tested statistically.

1. From 10 question instruments, two factors are obviously identifiable, measured by 5 questions of which they are comprised.
2. There is a significant prediction of student teachers’ speaking anxiety by gender, proficiency, and class type.

\textbf{METHOD}

\textbf{Participants and Setting}

The participants in this quantitative study were 72 second-year EFL students studying an English Teacher Training programme at a public university in Jambi, Indonesia. The questionnaire was initially distributed to all second-year students (n=102), but some of them seemed reluctant to return the questionnaire. The reason for selecting second-year EFL learners as respondents was because they had taken all speaking modules offered through their programme. Among the 72 respondents, 18 (25\%) participants were male, and 54 (75\%) were
female. 34 (47.2%) participants had previously taken English courses, while 38 (52.8%) had never taken any. 45 (62.5%) participants enrolled in regular classes, and 27 (37.5%) enrolled in regular-\textit{mandiri} classes. Table 1 details a summary of the sample.

<table>
<thead>
<tr>
<th>Table 1 Sample Distribution</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
</tr>
<tr>
<td>Male</td>
</tr>
<tr>
<td>Female</td>
</tr>
<tr>
<td><strong>Experience in English Course(s)</strong></td>
</tr>
<tr>
<td>Yes (at least once)</td>
</tr>
<tr>
<td>No (Never)</td>
</tr>
<tr>
<td><strong>Class Type</strong></td>
</tr>
<tr>
<td>Regular</td>
</tr>
<tr>
<td>Regular-\textit{Mandiri}</td>
</tr>
</tbody>
</table>

**Research Instruments**

For the purpose of gathering data from the participants, a close-ended questionnaire was used as a research instrument for this study. The statements in the questionnaire were constructed by developing Syakur’s (1987) theory on five speaking aspects, and adapted some elements of Horwitz’s (1986) questionnaire on FLCAS. The 13-item questionnaire was developed in English (see Appendix; questionnaire). The first three items related to the participants’ personal information: gender, English course experience and class type. The purpose of items 4-13 in the questionnaires was to explore the participants’ level of anxiety in the component of speaking: Q4 & Q5 (comprehension); Q6 & Q7 (grammar); Q8 & Q9 (word choice); Q10 & Q11 (pronunciation); Q12 & Q13 (fluency). The five point Likert scale was used in constructing statements’ options: (1) strongly disagree, (2) disagree, (3) neither agree nor disagree, (4) agree, and (5) strongly agree.
Once the questionnaire had been constructed, the researcher submitted it to an expert, the module tutor, to assess the validity of its content. Carmines and Zeller, cited in Cohen et al (2011) underline that content validity focuses on how “the instrument must show that it fairly and comprehensively covers the domain or items that it purports to cover” (p.188). Content validity is evidently fundamental in constructing questionnaires, aiming to determine whether the questionnaire measures all facets of a given construct or not. In this case, some revision to the questionnaire was required, primarily relating to sentence structures and wording.

**Data Collection and Data Analysis**

Following revision, the questionnaire was subsequently administered to the second year trainee teachers of an English Teacher Training programme at a public university in Jambi, Indonesia with the assistance of the researcher’s colleagues, specifically speaking module tutors. The researcher electronically sent a set of questionnaires to his colleagues and provided instructions around how the questionnaire should be correctly administered. Each tutor then administered the questionnaire to their class and returned the completed questionnaire electronically to the researcher.

When data collection had been completed, the researcher then manually input the data in Statistical Package for Social Sciences 22.0 (SPSS) and started quantitative analysis. The quantitative data was analysed using descriptive and inferential statistics, including explanatory factor analysis (EFA) and multiple linear regression analysis. EFA was applied to determine the construct structure while multiple linear regression analysis was applied to investigate anxiety levels and associated factors: gender, proficiency and class type. The findings were presented in tables and illustrated in figures.

To satisfy research ethical obligations, prior to undertaking the study, the researcher sent a letter requesting consent to the Director of the English Teacher Training programme. When permission had been granted from the research site, the researcher administered a consent
form, containing information of the nature and purpose of the study, to the prospective participants along with a copy of the questionnaire. They could decline the researcher’s invitation to take part in the study by not signing the form and/or returning the questionnaire. The signed consent form was used as proof that participants had voluntarily agree to take part and is documented for scholarly purpose only. To further protect the participants’ identity, anonymity was used in the final report.

**FINDINGS AND DISCUSSIONS**

As indicated, two hypotheses had been formulated in this study. In the first hypothesis, the researcher wanted to explore the factor structure underlying speaking anxiety responses in the data set. The second hypothesis was intended to investigate if the IVs - gender, proficiency, and class type- become the significant predictors towards learners’ speaking anxiety as a dependent variable (DV). Two inferential statistics, namely explanatory factor analysis and multiple linear regression, were used.

**Result of Factor Analysis**

The first step in conducting factor analysis is to produce the correlation matrix intended to determine whether or not the study variable correlates and, if so, the extent of the correlation. Tabachnick and Fidell (2001) outline that the use of factor analysis is questionable if a correlation, as reflected in the correlation matrix, is less than .30. Therefore, since the correlations yielded in the correlation matrix in this test exceeded .30, factor analysis was an appropriate data test. The result of analysis as suggests that all 10 variables are correlated, which indicates that there is a patterned relationship among the variables.

KMO and Bartlett’s test was utilised to confirm if the variables have a patterned relationship and to measure the sampling adequacy for the variables (Hair et al, 2014; Tabachnick and Fidell, 2013). The result of KMO and Bartlett’s test is detailed in the following table:
As shown in table 2, $p$ value of Bartlett’s test of Sphericity result is $< 0.001$ which means very significant. The significance of the result can be determined when $p$ value $\leq 0.05$ (Field, 2013; Hair et al, 2014; Tabachnick and Fidell, 2013). The small $p$ value shows that there is a statistically significant interrelationship between variables. The KMO measure of sampling adequacy (MSA) result is 0.881. This result exceeds the minimum cut-off point, above 0.50 (Tabachnick and Fidell, 2013; Hair et al, 2014), indicating that the data is sufficient for EFA.

Table 2 KMO and Bartlett's Test

| Kaiser-Meyer-Olkin Measure of Sampling Adequacy. | .881 |
| Bartlett's Test of Sphericity | Approx. Chi-Square | 282.579 |
|                           | df | 45 |
|                           | Sig. | .000 |

Table 3 Total Variance Explained

<table>
<thead>
<tr>
<th>Component</th>
<th>Initial Eigenvalues</th>
<th>Extraction Sums of Squared Loadings</th>
<th>Rotation Sums of Squared Loadings</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Total % of Variance</td>
<td>Cumulative %</td>
<td>Total % of Variance</td>
</tr>
<tr>
<td>2</td>
<td>1.030</td>
<td>10.296</td>
<td>58.488</td>
</tr>
<tr>
<td>3</td>
<td>.890</td>
<td>8.903</td>
<td>67.391</td>
</tr>
<tr>
<td>4</td>
<td>.746</td>
<td>7.461</td>
<td>74.853</td>
</tr>
<tr>
<td>5</td>
<td>.574</td>
<td>5.743</td>
<td>80.595</td>
</tr>
<tr>
<td>6</td>
<td>.505</td>
<td>5.052</td>
<td>85.647</td>
</tr>
<tr>
<td>7</td>
<td>.444</td>
<td>4.443</td>
<td>90.091</td>
</tr>
<tr>
<td>8</td>
<td>.392</td>
<td>3.922</td>
<td>94.013</td>
</tr>
<tr>
<td>9</td>
<td>.346</td>
<td>3.456</td>
<td>97.468</td>
</tr>
<tr>
<td>10</td>
<td>.253</td>
<td>2.532</td>
<td>100.000</td>
</tr>
</tbody>
</table>

Extraction Method: Principal Component Analysis.

Table 3 shows the total variance explained and indicates the number of significant factors. Kaiser’s rule, at this stage, was used to determine the most eligible factors for interpretation by extracting the
factors from the variables data. Kaiser’s criterion (Kaiser, 1958) stipulates that the only components with an eigenvalue greater than or equal to 1.0 should be retained for further analysis.

Table 3 clearly indicates that there are only two significant factors from the data. Together they are capable of explaining roughly 58.5% of all the variable variances. In addition, figure 1 below confirmed the findings of retaining 2 factors.

![Scree Plot](image)

Figure 1 Screen Plot

The two components with eigenvalue greater than 1.0 were rotated using Varimax rotation technique (Comrey and Lee, 2009; Hair et al, 2014; Stevens, 2009; and Tabachnick and Fidell, 2001) to generate rotated component matrix. Tabachnick and Fidell (2001) state that a high value of variable loading, reflected in rotation component matrix, shows the pure measurement of the component. Hair et al (2014) comment that each individual variable loading value, reflected in rotated component matrix, should be .50 or greater to provide interpretive value and to indicate the interrelation of the variables in the factor.

By looking at the minimum loading value, as proposed by Hair et al (2014), it is obvious which variables belong to each component. As shown in table 4, component 1 consists of 7 variables with .81 as the highest factor weight value and .59 as the lowest value. Component 2, on the other hand, comprises 3 variables with the sequence loading values: .82, .72, and .61. These values in each
component imply that variables are highly interrelated. Nevertheless, it seems impossible for the researcher to label the factors as there are some overlapping variables in each component. In this respect, a factor analysis yielded an unexpected result showing different items are loading in different subscales. Of the 7 variables in component 1, for instance, 4 items are apprehension subscale and the other 3 items belong to feedback subscale. Such findings suggest that a factor analysis does not really support the items that are supposed to be in a particular subscale and the hypothesis is therefore rejected.

The unexpected result of factor analysis is likely to occur for several reasons. Firstly, the sample of research might not be sufficient because only 72 participants were involved in this study. A wide range of recommendations in relation to sample in factor analysis has been made. Experts have different opinions regarding the minimum number of subjects required for analysis: 100 (Gorsuch, 1983; Kline, 1979); 150 (Hutcheson and Sofroniou, 1999); 200 (Guilford, 1954); 250 (Cattell, 1978); 300 (Nunnally, 1994; Tabachnick and Fidell, 2001, 2013). Comrey and Lee (1992) provided the following guidance regarding the adequacy of sample size: 100 = poor, 200 = fair, 300 = good, 500 = very good, 1,000 or more = excellent.

Table 4 Rotated Component Matrixa

<table>
<thead>
<tr>
<th>Component</th>
<th>1</th>
<th>2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feedback Pronunciation (11)</td>
<td>.811</td>
<td>.136</td>
</tr>
<tr>
<td>Feedback Grammar (7)</td>
<td>.801</td>
<td>.154</td>
</tr>
<tr>
<td>Word Choice Apprehension (8)</td>
<td>.757</td>
<td>.252</td>
</tr>
<tr>
<td>Feedback Word Choice (9)</td>
<td>.705</td>
<td>.327</td>
</tr>
<tr>
<td>Grammar Apprehension (6)</td>
<td>.661</td>
<td></td>
</tr>
<tr>
<td>Fluency Apprehension (12)</td>
<td>.633</td>
<td>.413</td>
</tr>
<tr>
<td>Pronunciation Apprehension (10)</td>
<td>.596</td>
<td>.408</td>
</tr>
<tr>
<td>Feedback Comprehension (4)</td>
<td>.129</td>
<td>.816</td>
</tr>
<tr>
<td>Feedback Fluency (13)</td>
<td>.365</td>
<td>.716</td>
</tr>
<tr>
<td>Comprehension Apprehension (5)</td>
<td>.153</td>
<td>.604</td>
</tr>
</tbody>
</table>

The unexpected result of factor analysis is likely to occur for several reasons. Firstly, the sample of research might not be sufficient because only 72 participants were involved in this study. A wide range of recommendations in relation to sample in factor analysis has been made. Experts have different opinions regarding the minimum number of subjects required for analysis: 100 (Gorsuch, 1983; Kline, 1979); 150 (Hutcheson and Sofroniou, 1999); 200 (Guilford, 1954); 250 (Cattell, 1978); 300 (Nunnally, 1994; Tabachnick and Fidell, 2001, 2013). Comrey and Lee (1992) provided the following guidance regarding the adequacy of sample size: 100 = poor, 200 = fair, 300 = good, 500 = very good, 1,000 or more = excellent.
They further encourage researchers to obtain a minimum of 500 samples whenever possible. Adequate sample size is therefore necessary because it produces a more accurate solution (Costello and Osborne, 2005). Secondly, content validity of the questionnaire may be another concern. As previously stated, this type of validity mainly focuses on how the questionnaire addresses all facets of phenomena (Carmines and Zeller in Cohen et al, 2011). The total number of items seems inadequate to measure all facets because there were only 10, that is, 5 items for each subscale with every item representing the component of speaking category. Additional items in subscale and/or categories may be best to the meaningful factorial solution.

<table>
<thead>
<tr>
<th></th>
<th>Cronbach's Alpha</th>
<th>Cronbach's Alpha Based on Standardized Items</th>
<th>N of Items</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>.876</td>
<td>.876</td>
<td>10</td>
</tr>
</tbody>
</table>

Although this questionnaire has issues concerning its construct validity, it is still found to be reliable. This is because the Cronbach alpha value is .876 which is more than its acceptable value .70 (Hair et al, 2014; Nunnally and Bernstein, 1978) as shown in the table 5. Thus, further analysis can be conducted using this questionnaire.

**Result of Multiple Linear Regression**

Multiple linear regression is generally defined as one of the modelling techniques that enable a researcher to assess the relationship between a DV (predicted) and some IVs (predictor). The final result of this modelling is the development of a regression equation (line of best fit) between DV and some IVs. Prior to conducting regression analysis, the researcher initially recoded IVs (gender, proficiency and class type) into dummy variables as they are nominal variables. They are dummy coded by assigning ‘0’ to one category and ‘1’ to another category and respectively named as: ‘female’ (‘0’ → male, ‘1’ → female); ‘proficient’ (‘0’ → not having
course experience, ‘1’ → having course experience); ‘regular’ (‘0’ → non regular class, ‘1’ → regular class). The researcher then tested some assumptions, including multicollinearity, data normality, linearity, and outliers.

- Multicollinearity

Multicollinearity is a phenomenon in which more than one of the IVs are highly correlated in multiple regression (Field, 2013; Hair et al, 2014; Stevens, 2009). A researcher must check this assumption to ensure no predictor variables are correlated in order to avoid the replication of the tendency predictor variable. The result of the correlation test, in table 6, clearly demonstrates that there is no existence of multicollinearity case in this study’s data-set. No correlation is found between predictor variables as all the corresponding values are less than .70.

<table>
<thead>
<tr>
<th></th>
<th>Total anxiety scale</th>
<th>Female</th>
<th>Proficient</th>
<th>Regular</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pearson</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Correlation</td>
<td>1.000</td>
<td>-.115</td>
<td>-.356</td>
<td>-.171</td>
</tr>
<tr>
<td>Female</td>
<td>-.115</td>
<td>1.000</td>
<td>.161</td>
<td>-.116</td>
</tr>
<tr>
<td>Proficient</td>
<td>-.356</td>
<td>.161</td>
<td>1.000</td>
<td>.215</td>
</tr>
<tr>
<td>Regular</td>
<td>-.171</td>
<td>-1.16</td>
<td>.215</td>
<td>1.000</td>
</tr>
<tr>
<td>Sig. (1-tailed)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>scale</td>
<td></td>
<td>.167</td>
<td>.001</td>
<td>.075</td>
</tr>
<tr>
<td>Female</td>
<td>.167</td>
<td>.</td>
<td>.089</td>
<td>.166</td>
</tr>
<tr>
<td>Proficient</td>
<td>.001</td>
<td>.089</td>
<td></td>
<td>.035</td>
</tr>
<tr>
<td>Regular</td>
<td>.075</td>
<td>.166</td>
<td>.035</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total anxiety</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>scale</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Female</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>Proficient</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
</tr>
<tr>
<td>Regular</td>
<td>72</td>
<td>72</td>
<td>72</td>
<td>72</td>
</tr>
</tbody>
</table>
• Data Normality, Linearity, and outliers.

The researcher is advised to check the data normality and linearity before reporting the model of regression analysis. The normality and linearity of data distribution can be seen from the normal P-P plot regression. If the dots, reflected in P-P plot, are perfectly straight-lined, it can be inferred that the data are highly normally distributed. The finding shows that the data are normally distributed because the dots are reasonably close to the best fit line as shown in chart 2. There were some deviations from the perfect line, but they were not major ones.

![Normal P-P Plot of Regression Standardized Residual](image_url)

**Figure 2 Normal P-P Plot**

Aside from the normal distribution shown in figure 2 above, the researcher also provides some evidence of normality data from other measurements, including skewness, kurtosis, Kolmogorov–Smirnov test, and boxplot. In an attempt to check the possible outliers in the data, scatterplot produced by SPSS was further analysed. If the dots are clustered between the threshold -3 and 3, it indicates that the data is acceptable and there are no outliers. The findings of this study, as shown in figure 3, suggest that there are no outliers found in the data. Although the dots are not clustered closely to each other, they are still scattered in the specified threshold.
After assumptions are met, the next stage is to evaluate the model, the aim of which is to determine the model effectiveness, significance, and its accurate predictor(s). In order to evaluate the model, the first consideration should be the ANOVA table, which is produced by SPSS.

Table 7 ANOVA

<table>
<thead>
<tr>
<th>Model</th>
<th>Sum of Squares</th>
<th>df</th>
<th>Mean Square</th>
<th>F</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression</td>
<td>493.552</td>
<td>3</td>
<td>164.517</td>
<td>3.743</td>
<td>.015</td>
</tr>
<tr>
<td>Residual</td>
<td>2988.448</td>
<td>68</td>
<td>43.948</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>3482.000</td>
<td>71</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 7 above reports the results of an ANOVA test that represents an overall test value of how well the model as a whole fits the data. Hair et al (2014) and Steven (2009) underscore that the model is deemed statistically significant to predict the variation in DV if p-value is ≤0.05. The significant value (.015), as can be seen in the table 7, shows that there is a strong evidence that the model is statistically significant and useful to predict the variation in the predicted variable. Therefore, the alternative hypothesis of this study is accepted.

In order to see how much of the variability in the outcome is accounted for by the predictors, the value of adjusted $R^2$ in the table
Abrar, An Investigation into Indonesian EFL University Students’ Speaking Anxiety

of ‘model summary’ should be considered. Hair et al (2014) posit that adjusted \( R^2 \) is a “modified measure of the coefficient and determination that takes into account the number of independent variables included in the regression equation and the sample size” (pp.152) and is regarded as the best estimate of the degree of the relationship in the basic population. The following table shows the value of adjusted \( R^2 \) of this model.

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R Square</th>
<th>Adjusted R Square</th>
<th>Std. Error of the Estimate</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.376</td>
<td>.142</td>
<td>.104</td>
<td>6.629</td>
</tr>
</tbody>
</table>

The value of adjusted \( R^2 \) of this study is fairly small, .104, which means that the model, using 3 predictor variables, only explains 10.4% of the variance in speaking anxiety. This relatively small adjusted \( R^2 \) value also implies that nearly 90% of the variance in speaking anxiety is explained by other factors. This unexpected result encouraged the researcher to see its computed size effect. By running G* power software, it is found that effect size \( f^2 \) is .116. This indicates that the computed effect size corresponds to a small to medium effect size. To sum up, this model is statistically significant (\( R^2= .104, F(3, 68)= 3.743, p< .05, f^2\text{cohen}= .116 \)).

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized Coefficients</th>
<th>Standardized Coefficients</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>B</td>
<td>Std. Error</td>
<td>Beta</td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>(Constant)</td>
<td>34.527</td>
<td>1.971</td>
<td>17.520</td>
</tr>
<tr>
<td></td>
<td>Female</td>
<td>-1.234</td>
<td>1.851</td>
<td>-.077</td>
</tr>
<tr>
<td></td>
<td>Proficient</td>
<td>-4.454</td>
<td>1.633</td>
<td>-.320</td>
</tr>
<tr>
<td></td>
<td>Regular</td>
<td>-1.597</td>
<td>1.673</td>
<td>-.111</td>
</tr>
</tbody>
</table>

Table 9 provides information regarding the relative strength of individual predictor variables. Only the proficient variable, of the three
predictors, seems to contribute more to the model as its p-value was found significant (p<.05), while female and regular variables do not contribute (p_{female}>.05 and p_{regular}>.05). Another way to see the significant predictor is from the standardized coefficient column (Beta-value). The proficient variable appears to have a greater effect (-.320) to the model than the other variables. Thus, on the basis of the mentioned values, it is tempting to conclude that the only useful predictor is the proficient variable. Table 9 further represents the information needed to construct the actual model. Using the ‘unstandardized coefficient’ from the table, the construct of statistical model can be formulated as follows:

\[
\text{Predicted ‘Speaking Anxiety’} = 34.572 - 1.234 \times \text{‘Female’} - 4.454 \times \text{‘Proficient’} - 1.597 \times \text{‘Regular’}.
\]

The regression equation above shows that if ‘female’ is increased by one unit, total speaking anxiety decreases 1.234 units. In other words, female students are found to experience less anxiety than male students. This result surprisingly contradicts the majority of research (Çağatay, 2015; Park and French, 2013; Ekström, 2013; Öztürk, and Gürbüz, 2013; Hsu, 2012; Tianjian, 2010; Occhipinti, 2009; Huang, 2004) in finding that female students tended to be more anxious at speaking foreign language than male students. However, from a second language perspective, the result of this study seems reasonable in that both genders, by nature, are different in the process of acquiring a language. Females are considered to be more motivated in learning a second language than males (Gardner and Lambert, 1972).

With regard to ‘proficient’ predictor, the model suggests that the total speaking anxiety significantly decreases as much as 4.454 units if ‘proficient’ is increased by one unit. It simply shows that proficient learners are less nervous about speaking than non-proficient learners. This result is congruent to the findings of other studies (Çağatay, 2015; Tercan and Dikilitas, 2015; Zhao and Whitchurch, 2011; Tianjian, 2010; Liu, 2006). This study’s finding is
rationally accepted as proficient-language learners have not only learned the language at school, but also at informal institution(s). They possibly gain more experiences at speaking practice and therefore their anxiety level gradually decreases. Skehan (1989) argues that “students at higher levels might enjoy wider repertoire of behaviours which would help them to deal with anxiety in language learning contexts more flexibly” (p.116). In addition, the model indicates that regular-mandiri students feel more anxious than regular students as if ‘regular’ is increased by one unit, speaking anxiety generally decreases 1.594 units.

CONCLUSION
This study aims to identify the constructs of a speaking anxiety questionnaire and to explore the model for predicting speaking anxiety based on gender, proficiency and class type. Firstly, the latent constructs of speaking anxiety responses in the dataset are determined through an explanatory factor analysis. The findings reveal two underlying factors. Nevertheless, these two factors are difficult to name due to some overlapping variables in each component. Therefore, additional participants and questionnaire items may be included for producing more accurate and meaningful factorial solution. Secondly, this study aims to examine the relationship between predictors (gender, proficiency, and class type) and a predicted variable (speaking anxiety) with a multiple linear regression test. The findings suggest that the resultant regression model is determined to be statistically significant for predicting the variation in speaking anxiety. In addition, among the three predictors, the proficient variable appears to be the most significant predictor in that it contributes more to the model.

REFERENCES


**Abrar**, An Investigation into Indonesian EFL University Students’ Speaking Anxiety


Abrar, An Investigation into Indonesian EFL University Students’ Speaking Anxiety


Abrar, An Investigation into Indonesian EFL University Students' Speaking Anxiety


Abrar, An Investigation into Indonesian EFL University Students’ Speaking Anxiety


