EFL Learners’ Foreign Language Learning Anxiety and Language Performance: A Meta-Analysis Study

Melih Dikmen

1Fırat University, ORCID ID: 0000-0001-7015-6236

To cite this article:

EFL Learners’ Foreign Language Learning Anxiety and Language Performance: A Meta-Analysis Study

Melih Dikmen

1Fırat University

Abstract

The relationship between foreign language anxiety (FLA) and English as a foreign language (EFL) learners’ performance is constantly being examined through experimental and theoretical studies. The aim of this study is to examine the association between FLA and EFL performance through meta-analysis. As a result of an extensive literature review, 69 studies from fourteen countries (N= 23,150) were identified as eligible to be included in the analysis process. The overall correlation between FLA and EFL performance was calculated as -.61. In terms of moderator variables, the results revealed that while grade level and country in which studies were conducted did not have an effect on the association between FLA and EFL performance, type of anxiety was found to have a moderator effect. It was observed that listening, test, speaking, reading, and writing anxiety affected EFL learning performance respectively from the most to the least. Finally, publication year and sample size were found to have a significant effect on the association between FLA and EFL performance through meta-regression analysis. Based on the findings, it is concluded that in order to increase EFL learners’ performances, their foreign language anxiety should be decreased.

Key words: EFL, Foreign language learning anxiety, Language performance, Meta-analysis

Introduction

Today, many people spend time to learn English as a foreign language (EFL). Due to the complex and multi-dimensional nature of learning foreign language, people may experience some difficulties in the learning process. In the last decade, many studies were conducted in order to determine the factors that negatively influence language learning process (Çakıcı, 2016). Anxiety is one of the affective variables that researchers mainly examined in foreign language learning (Teimouri, Goetze & Plonsky, 2019). Anxiety is defined as a state of arousal that an individual experiences when faced with a stimulus that causes physical, emotional and/or mental changes (Cüceloğlu, 2005). Foreign language anxiety is observed in situations in which formal learning of foreign language occurs, especially when an individual has low communication abilities in that language (Rodríguez & Abreu, 2003). Yoshida (2010) reported that students attend foreign language classes with anxiety. The fact that the etiology of FLA is not fully known has increased the number of studies on this subject. In one of the first studies, Cattell and Scheier (1961) stated that second language (L2) learning anxiety could be caused by personality traits or a temporary emotional state manifested at a certain time. However, these early studies were defined as the confounded approach by MacIntyre (2017) since those studies revealed inconsistent results due to the limitations in considering the meaning of anxiety for language learners in detail. On the other hand, Tsiplakides and Keramida (2009) found that students with anxiety avoid participating in-class activities since they believe that they have ineffective speaking skills and they are concerned that their classmates may criticize them due to their speaking skills. Moreover, anxiety is observed when students are lack of self-confidence, compare themselves with others, and have negative thoughts about foreign languages (Öztürk, 2003). MacIntyre (2017) argued that FLA is affected by internal physiological processes, cognitive and emotional states, and the presence of other people. Besides, one of the main triggers of FLA is the fear of other’s negative evaluation (Aydın, 2008). In addition, studies indicate that FLA may be caused by many internal (Cheng, 2002; Liu & Jackson, 2008; Mak, 2011; Dewaele, 2017) and external (Liu, 2006; Mak, 2011; Oxford, 2017) factors.

* Corresponding Author: Melih Dikmen, melihdikmen@gmail.com
FLA can be extremely harmful to the language learning process (MacIntyre, 2017). FLA is negatively related to students’ cognitive learning process and their interaction and communication skills. In addition, FLA causes student difficulty in focusing on course content during lessons (Aguila & Harjanto, 2016). Also, students with anxiety experience a mental block when performing an in-class activity (Chen & Chang, 2004). It was also reported that the higher the FLA level of the students, the lower their verbal performance scores (Azizifar, Faryadian & Gowhary, 2014). FLA negatively affects learners’ EFL performance by interfering with the EFL learning process (Liu & Xiangming, 2019; Kabigting & Nanud, 2020). In addition, studies revealed that unlike other types of anxiety, FLA causes students to develop the fear of being evaluated negatively, which, in turn, negatively affects the language learning process (Cheng, Horwitz & Schallert, 1999).

Many studies uncovered the negative association between FLA and EFL performance (Horwitz, 2001; Elkhafaifi, 2005; Birjandi & Alemi, 2010; Ghorbandordinejad & Ahmadosad, 2016; Liu & Xiangming, 2019; Kabigting & Nanud, 2020). More specifically, some studies argued that a higher FLA level negatively affects EFL performance (Hewitt & Stephenson, 2011; Rezaabadi, 2016; Elmai-Ozssaray & Eren, 2018; Yurtseven & Akpur, 2018). However, little is known about how FLA is affected by and relates to other learner traits (Dörnyei, 2005; Dewaele, 2013; Dewaele & Ip, 2013). Despite the discussions about FLA, some researchers rejected the concept of FLA (Javanbakht & Hadian, 2014; Lian & Budin, 2014; Sparks, Ganschow & Javorsky, 2000; Tridinanti, 2018). Specifically, Sparks and colleagues (2000) stated that the anxiety experienced by language learners is due to potential learning difficulties and even limitations in their first language.

No meta-analysis study was found in the literature to explain the contradictory findings regarding the relationship between FLA and EFL performance and to examine the variables that may be effective in the relationship between FLA and EFL. However, there are meta-analyses studies on the relationship between L2 learning and FLA (Teimouri, Goetze & Plonsky, 2019; Zhang, 2019). Teimouri, Goetze, and Plonsky (2019) reported that the overall correlation between FLA and L2 learning was −.36 (95% CI [.39, −.33]). In the same study, although moderating effects of education level, target language, and anxiety were examined, the exact effects of the moderators were not reported. Therefore, no general conclusion can be drawn about the significance of the moderators. Zhang (2019) also conducted a meta-analysis on the relationship between L2 learning and FLA. Unlike Teimouri, Goetze, and Plonsky’s (2019) meta-analysis study, Zhang did not include the grade point averages (GPA) reported by the students in the meta-analysis. The rationale for this exclusion was the reliability issues in the reported GPAs by students. Also, Zhang (2019) included Chinese, English, Persian, Arabic, and Spanish languages and examined the effect of some moderator variables (anxiety type, age, proficiency, lexical similarity, language family, and publication year) on the relationship between L2 learning and FLA. The results revealed that anxiety type, lexical similarity, age, and publication year had an effect on the anxiety-L2 learning relationship. Also, it was found that proficiency, language family, and publication type do not have a significant modulating effect.

The fact that English is a common language used for communication in various fields in the world has increased the importance of examining the factors that affect the performance of EFL learners. The meta-analysis studies on EFL in the literature were unable to provide satisfactory findings about the relationship between EFL and FLA. Also, in addition to the moderators (education level and publication year) examined in the previous meta-analysis studies, a number of additional moderators including country, type of anxiety, and sample size may be critical to examine the relationship between L2 learning and FLA. A number of criteria were taken into account in determining the moderators of the current study. The rationale for the inclusion of those variables is provided below:

- The previous meta-analysis studies (Teimouri, Goetze & Plonsky, 2019) reached a limited number of studies in terms of elementary and middle school levels. Also, it was suggested that future studies should consider inclusion of various education levels. Thus, the education level was included as a moderator variable in the present study.

- For EFL learners, it may be important to determine how language anxiety relates to language learning performance by time. For this reason, publication year was included as a moderator variable in the present study.

- In the previous studies (Zhang, 2019), language family was considered as a moderator variable on the relationship between L2 learning and FLA and its non-significant effect was reported. On the other hand, Yoşurtçu and Yoşurtçu (2013) stated that culture has a critical role on anxiety. Thus, country where studies were conducted was included as a moderator variable in the present study.

- Language learning includes reading, writing, listening, and speaking skills. Melchor-Couto (2017) stated that these skills in language learning may trigger FLA. Thus, type of anxiety (reading, writing, listening, speaking, and test anxiety) was included as a moderator variable in the present study.
• Tuncer and Dikmen (2017) discussed the dilemma about whether the standard error increases in line with the sample size. Thus, the sample size was included as a moderator variable in the present study.

The Current Study

This study aimed to examine the direction (positive or negative) and magnitude of the relationship between FLA and EFL performance. Beyond the results of the meta-analysis in the literature, this study also aimed to determine some variables’ moderating effects on the relationship between FLA and EFL performance. Thus, the following research questions were addressed:

1. What is the direction and magnitude of the relationship between FLA and EFL performance?
2. Do grade level, country, type of anxiety, publication year, and sample size have a significant moderating effect on the relationship between FLA and EFL performance?

Method

The goal of the present meta-analysis is to systematically examine the relationship between FLA and EFL performance and to gather the findings of the previous studies to provide a holistic perspective. Meta-analysis enables researchers to combine findings of the studies, to synthesize, and to calculate the overall effect size (Lipsey & Wilson, 2000; Card, 2012). In addition, several analyses were conducted for certain variables that may have a potential to influence EFL performance.

Moderating Variables in the Study

The moderating variables are the variables that may influence the observed effect size. In this particular study, the following variables are considered as the moderating variables: publication year (between 2010 and 2020), grade level (middle school, high school, and higher education level), country (Turkey, China, Iran, India, Indonesia, Korea, Malaysia, Pakistan, Philippines, Poland, Saudi Arabia, Spain, Taiwan and Yemen), and type of anxiety (reading, writing, listening, speaking, and test anxiety).

Data Sources and Search Strategies

In order to identify eligible studies, ten electronic databases were searched, including Web of Science, ERIC (EBSCO), Scopus (A&I), ULAKBIM, Taylor & Francis Online, Science Direct, Springer LINK, Wiley Online Library Full Collection, Google Scholar, ProQuest Dissertation, and Turkey Council of Higher Education Thesis Center. In order to conduct the search, the following keywords were used: (i) foreign language anxiety, (ii) foreign language anxiety and academic achievement, (iii) foreign language anxiety and academic performance. The search was conducted on December 30, 2020 by using the keywords through the databases and only the studies written in English or Turkish were selected.

Inclusion and Exclusion Criteria

The following criteria were followed in order to determine eligible studies: (i) studies must be published between 2010 and 2020, (ii) studies must be published in either Higher Education Thesis Center or peer-reviewed journals, (iii) studies must include sufficient amount of statistical information (sample size and correlation values), (iv) full-text documents must be reached, and (v) studies must be written in either Turkish or English. Excluded studies usually were out of scope, included qualitative data, did not have full-text documents, and/or did not have sufficient statistical information. Also, if an author had a thesis and an article derived from the thesis, only one of them was included in the analysis process.

After the search in the databases, the titles and abstracts were examined to ensure they were within the scope of this particular study. Then, full-text documents were retrieved. The PRISMA guidelines (Moher, Liberati, Tetzlaff, Altman, & Prisma Group, 2009) showing the acquisition process of the eligible studies is given in Figure 1.
As seen in Figure 1, a total of 3022 studies were identified. After reviewing those studies, 351 studies were excluded due to duplication. Also, 2116 studies were also excluded because of their scope, which left a total of 755 studies. Based on the inclusion criteria, 690 studies were omitted. As a result, 69 studies were included in the meta-analysis process. There were 13 studies from Turkey, 20 studies from China, 18 studies from Iran, 2 studies from Yemen, 3 studies from Indonesia, Malaysia and Saudi Arabia, and 1 from India, Korea, Pakistan, Poland, Philippines, Spain and Taiwan.

Coding Method

One of the critical steps in meta-analysis is data coding in order to combine or compare the results of the eligible studies. Therefore, an appropriate coding form was designed to compare the studies. The coding form is provided in Table 1.

<table>
<thead>
<tr>
<th>Study ID</th>
<th>Content of the Study</th>
<th>Data of the Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Title</td>
<td>Country where conducted</td>
<td>Sample size (N)</td>
</tr>
<tr>
<td>Author(s)</td>
<td>Grade level</td>
<td>Correlation (r)</td>
</tr>
<tr>
<td>Publication year</td>
<td>Type of anxiety</td>
<td></td>
</tr>
<tr>
<td>Publication type</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to increase the reliability of the present study, each eligible study was independently coded by two researchers. The coders had doctorate degree and had expertise in the field of qualitative research. After individual coding, the experts discussed about the differences in their coding and disagreements were solved with consensus. The inter-observer agreement value was calculated as .92. This value shows that there is a perfect match between the coders (Viera & Garreth, 2005).
Data Analysis and Interpretation

In meta-analysis studies, the effect size constitutes the basis of the research. There are three different models in the literature: fixed effect, random effect, and mixed effect. The fixed effect model accepts that the studies included in the analysis are homogeneous and that the differences in effect size are caused by sampling errors. The random effect model also agrees that the studies included in the analysis are homogeneous and that the differences in effect size are caused by sampling errors and the characteristics of the studies examined (Cooper, 2010). On the other hand, in the mixed effect model, it is assumed that the determined differences in effect size may be due to sampling errors, differences between studies and random factors (Borenstein, Hedges, Higgins, & Rothstein, 2010). In meta-analysis studies, researchers should examine the heterogeneity when deciding which model to adapt (Borenstein, Hedges, Higgins, & Rothstein, 2010). Borenstein and colleagues (2013) suggest use of chi-square or similar tests in order to determine the heterogeneity. If the Q value obtained from the heterogeneity test is less than the Q value shown in the χ² table, the homogeneity status met, and if it is greater, the heterogeneity status is fulfilled (Dinçer, 2014). The significance of the Q statistics calculated in this test (p < .05) suggests that the studies are heterogeneous. In addition, considering that the studies included in the present study were conducted in different countries and with different sample sizes may cause a difference in the calculated effect sizes, the findings in this study were interpreted according to the random effect model. In order to calculate the effect size, Cohen’s d was used. Cohen’s d is an effect size used to indicate the standardized difference between two means. In addition, the following variables were used as moderator variable in analysis: grade level, anxiety type, country, sample size, and publication year. In meta-analysis studies, categorical moderators are examined through ANOVA and continuous moderators are examined through meta-regression analysis. Meta-regression is an extension to subgroup analyses that allows the effect of continuous characteristics to be investigated, and in principle allows the effects of multiple factors to be investigated simultaneously.

There exist some classifications in interpretation of the calculated effect sizes. The most widely used ones are Cohen’s (1977), Lipsey and Wilson’s (2001), and Thalheimer and Cook’s (2002) classifications. According to Cohen (1977), the effect size can be small if it is between .20 and .49, medium if it is between .50 and .79, and large if it is over .80. According to Lipsey and Wilson (2001), it is suggested that effect sizes of .15 are small, .45 are medium, and .90 are high. On the other hand, Thalheimer and Cook (2002) consider the effect size as negligible if -0.15 < d < 0.15, small if 0.15 < d < 0.40, medium if 0.40 < d < 0.75, large if 0.75 < d < 1.10, very large if 1.10 < d < 1.45, and huge if 1.45 < d. Interpretation of the effect sizes of the studies that were analyzed was conducted according to the classification of Cohen (1977).

Another critical aspect to be considered in meta-analysis studies is publication bias. Publication bias refers to situations in which studies with significant results are less likely to be published compared with studies with insignificant results. Therefore, it is highly possible to obtain high effect size (Borenstein, Cooper, Hedges & Valentine, 2009). In order to evaluate the presence of possible publication bias, funnel plots were drawn and Begg and Mazumdar rank correlation test was performed.

In order to combine the effect sizes obtained from the studies, the normal distribution of the effect sizes was tested. In this particular study, normal Q-Q plot and kurtosis and skewness coefficients were used to determine whether the effect sizes of the eligible studies had normal distribution. Normal Q-Q plot shows the association between the observed and expected values. When the observed values and expected values overlap, a 45-degree line emerges. The Q-Q graph regarding the distribution of the effect sizes of the 69 studies is given in Figure 2.
As seen in the figure, the effect sizes of the 69 studies cumulated along a line. However, this is not sufficient to decide whether the data is normally distributed (Can, 2013). Additionally, kurtosis and skewness values were calculated. The skewness (-.39) and kurtosis (1.47) values of the effect sizes were in the normal distribution range (-1.96 and +1.96) (Tabachnick & Fidell, 2007).

In this study, the Comprehensive Meta-Analysis v.2 (CMA v. 2.0) statistics package was used for effect sizes, heterogeneity test, moderator, meta-regression, and publication bias. Also, in order to calculate the inter-observer agreement value, the IBM SPSS 22.0 statistics package was used. A value of p < .05 was accepted as statistically significant.

Results

In this section, the findings related to the descriptive results, the calculated effect sizes, and the changes in sub-categories are provided.

Descriptive Results of the Studies

The eligible studies were examined in terms of publication year, type of publication, country where conducted, grade level, and type of anxiety. Descriptive results are provided in Table 2.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Frequency (f)</th>
<th>Percentage (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Publication Year</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2010/2011</td>
<td>5/7</td>
<td>7.25/10.14</td>
</tr>
<tr>
<td>2012/2013</td>
<td>2/10</td>
<td>2.90/14.49</td>
</tr>
<tr>
<td>2014/2015</td>
<td>9/6</td>
<td>13.04/8.70</td>
</tr>
<tr>
<td>2016/2017</td>
<td>6/8</td>
<td>8.70/11.59</td>
</tr>
<tr>
<td>2018/2019</td>
<td>7/5</td>
<td>10.14/7.25</td>
</tr>
<tr>
<td>2020</td>
<td>4</td>
<td>5.80</td>
</tr>
<tr>
<td><strong>Country</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>13</td>
<td>20.97</td>
</tr>
<tr>
<td>China</td>
<td>20</td>
<td>32.26</td>
</tr>
<tr>
<td>Iran</td>
<td>18</td>
<td>29.03</td>
</tr>
<tr>
<td>Indonesia</td>
<td>3</td>
<td>4.84</td>
</tr>
<tr>
<td>Malaysia</td>
<td>3</td>
<td>4.84</td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>3</td>
<td>4.84</td>
</tr>
<tr>
<td>Yemen</td>
<td>2</td>
<td>3.23</td>
</tr>
<tr>
<td><strong>Grade level</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>13</td>
<td>19.12</td>
</tr>
<tr>
<td>Higher education</td>
<td>55</td>
<td>80.88</td>
</tr>
<tr>
<td><strong>Type of anxiety</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking anxiety</td>
<td>12</td>
<td>17.39</td>
</tr>
<tr>
<td>Writing anxiety</td>
<td>3</td>
<td>4.35</td>
</tr>
<tr>
<td>Listening anxiety</td>
<td>10</td>
<td>14.49</td>
</tr>
<tr>
<td>Reading anxiety</td>
<td>5</td>
<td>7.25</td>
</tr>
<tr>
<td>Test anxiety</td>
<td>39</td>
<td>56.52</td>
</tr>
</tbody>
</table>

According to the results, it was observed that the studies included in the meta-analysis were conducted in 2013 at the most, while at least in 2020. In addition, the sample group was mainly higher education students. The studies were mainly conducted in China, Iran, and Turkey. In terms of type of anxiety, the studies generally focused on test anxiety. It was determined that the number of samples reached within the scope of 69 studies was 23.150.
Publication Bias Results

In order to determine possible publication bias, funnel plot was drawn and Begg and Mazumdar rank correlation test was performed. The funnel plot was provided in Figure 3.

As seen in the figure, the funnel plot does not present an asymmetric funnel, revealing that the eligible studies do not have publication bias. In order to ensure the absence of publication bias, Begg and Mazumdar rank correlation test was performed. The results are provided in Table 3.

Table 3. Begg and Mazumdar rank correlation results

<table>
<thead>
<tr>
<th>Kendall's S Statistics (P-Q)</th>
<th>-290.00000</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kendall’s tau value before the correction of continuity</td>
<td></td>
</tr>
<tr>
<td>Tau</td>
<td>-.12372</td>
</tr>
<tr>
<td>z value for Tau</td>
<td>1.50206</td>
</tr>
<tr>
<td>p (one-tailed)</td>
<td>.06654</td>
</tr>
<tr>
<td>p (two-tailed)</td>
<td>.13308</td>
</tr>
<tr>
<td>Kendall’s tau value after the correction of continuity</td>
<td></td>
</tr>
<tr>
<td>Tau</td>
<td>-.12329</td>
</tr>
<tr>
<td>z value for Tau</td>
<td>1.49688</td>
</tr>
<tr>
<td>p (one-tailed)</td>
<td>.06721</td>
</tr>
<tr>
<td>p (two-tailed)</td>
<td>.13442</td>
</tr>
</tbody>
</table>

According to the findings, Begg and Mazumdar rank correlation test did not produce significant result (tau b = -.12, p > .05). In other words, there was no publication bias in the study selection. Therefore, the analysis results were reliable.

Findings of General Effect Size

The studies examining the effect of FLA on language learning performance were gathered by using the random effect model. The results are provided in Table 4.

Table 4. Overall effect size, heterogeneity, and confidence intervals

<table>
<thead>
<tr>
<th>Model</th>
<th>N</th>
<th>Effect Size</th>
<th>Standard Error</th>
<th>Variance</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Z Value</th>
<th>P Value</th>
<th>Q Value</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Random</td>
<td>69</td>
<td>-.611</td>
<td>.049</td>
<td>.002</td>
<td>-.707</td>
<td>-.515</td>
<td>-12.473</td>
<td>.000</td>
<td>743.114</td>
<td>.000</td>
</tr>
</tbody>
</table>
The heterogeneity test produced a significant result (Q(model)=196.609, p≈.000). The overall effect size was found to be -0.61, which is a medium effect size as suggested by Cohen (1977). Therefore, it may be concluded that there was a medium and negative association between FLA and language learning performance. In addition, FLA explains 37% of the total variance of language learning performance. The forest plot showing the studies’ effect sizes and confidence intervals is provided in Figure 4.

![Forest plot](image_url)

**Figure 4.** The forest plow of random effects estimates for the included studies

The black square in the plow represents the effect size of the related study. The vertical horizontal lines indicate the 95% confidence interval of the effect size of the relevant study. According to the forest plot, it is seen that the study with the widest confidence interval and high weighted effect size belongs to Kabigting and Nanud (2020), while the smallest confidence interval and the lowest weighted effect size are in the study of Lan (2010). Also, the contribution of the other studies to the total effect size consists of weighted effect sizes close to each other.
Results of Moderator Analysis

There is a variation in the country, grade level, and type of anxiety variables among the studies. Thus, the effects of those variables on studies’ effect sizes were examined in order to determine their moderating effects. Since there was only one study conducted for some country and education level in terms of country and education level moderators, they were not included in meta-analysis. The findings are presented in Table 5.

Table 5. The effect sizes of studies on including moderators in relation to EFL performance

<table>
<thead>
<tr>
<th>Variables</th>
<th>Number of studies</th>
<th>Effect Size</th>
<th>Standard error</th>
<th>95% Confidence Interval</th>
<th>Qb</th>
<th>sd</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Grade level</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High school</td>
<td>13</td>
<td>-.495</td>
<td>.120</td>
<td>-.729 -.261</td>
<td>.949</td>
<td>1</td>
<td>.330</td>
</tr>
<tr>
<td>Higher education</td>
<td>55</td>
<td>-.623</td>
<td>.055</td>
<td>-.731 -.515</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>68</td>
<td>-.601</td>
<td>.050</td>
<td>-.699 -.503</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>China</td>
<td>20</td>
<td>-.480</td>
<td>.083</td>
<td>-.642 -.318</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Indonesia</td>
<td>3</td>
<td>-.084</td>
<td>.261</td>
<td>-.595 .427</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Iran</td>
<td>18</td>
<td>-.748</td>
<td>.096</td>
<td>-.936 -.560</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Malaysia</td>
<td>3</td>
<td>-.180</td>
<td>.239</td>
<td>-.650 .289</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>3</td>
<td>-.797</td>
<td>.233</td>
<td>-1.253 -.341</td>
<td>15.240</td>
<td>6</td>
<td>.018*</td>
</tr>
<tr>
<td>Turkey</td>
<td>13</td>
<td>-.708</td>
<td>.102</td>
<td>-.909 -.508</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yemeni</td>
<td>2</td>
<td>-.207</td>
<td>.273</td>
<td>-.742 .327</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>62</td>
<td>-.581</td>
<td>.049</td>
<td>-.677 -.485</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Type of anxiety</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Test anxiety</td>
<td>39</td>
<td>-.642</td>
<td>.066</td>
<td>-.771 -.513</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Listening anxiety</td>
<td>10</td>
<td>-.888</td>
<td>.132</td>
<td>-1.147 -.629</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Reading anxiety</td>
<td>5</td>
<td>-.443</td>
<td>.177</td>
<td>-.791 -.095</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Speaking anxiety</td>
<td>12</td>
<td>-.446</td>
<td>.122</td>
<td>-.685 -.206</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Writing anxiety</td>
<td>3</td>
<td>-.286</td>
<td>.221</td>
<td>-.719 .146</td>
<td>9.513</td>
<td>4</td>
<td>.049*</td>
</tr>
<tr>
<td>Total</td>
<td>69</td>
<td>-.611</td>
<td>.049</td>
<td>-.708 -.514</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*p < .05

In terms of grade level, FLA’s effects on EFL performance were -.495 for high schools, and -.623 for higher education. According to the findings, grade level did not have a significant moderating effect on the association between FLA and EFL performance (Qb = .949, p < .05).

According to the findings in Table 5, when considering the country as a moderator variable, the highest effect was observed in Saudi Arabia (d = -.797) and the lowest effect was observed in Indonesia (d = -.084). Indeed, country did have a significant moderating effect on the association between FLA and EFL performance (Qb = 15.240, p < .05).

In addition, type of anxiety was included in the analysis as a moderator variable. The results revealed that while listening anxiety had the highest effect (d = -.888) on EFL performance, writing anxiety had the smallest effect (d = -.286). Based on the findings, type of anxiety was a moderator variable that significantly influenced the association between FLA and EFL performance (Qb = 9.513, p < .05).

In meta-analysis studies, while categorical moderators are analyzed using the analog to the analysis of variance (ANOVA), continuous moderators are examined using multi-regression analysis. Since the year variable was a continuous variable, meta-regression analysis was performed. The findings are presented in Figure 5.
As seen in Figure 5, it is seen that there was a negative increase in the line slope as the publication year progressed from the past to the present. Table 6 provides the statistical results for this decrease.

Table 6. The statistical results for the publication year variable and the effect sizes

<table>
<thead>
<tr>
<th></th>
<th>Point Estimation</th>
<th>Standard Error</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Z value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>-0.02047</td>
<td>0.00488</td>
<td>-0.03004</td>
<td>-0.01090</td>
<td>-4.19383</td>
<td>0.00003*</td>
</tr>
<tr>
<td>Intercept</td>
<td>40.72116</td>
<td>9.83178</td>
<td>21.45122</td>
<td>138.44171</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The publication years of the studies from the past to the present caused a negative increase of 0.020 in the effect size, which is statistically significant (p<.05).

Another moderator variable in this study was the sample size of the studies. The meta-regression analysis results are presented in Figure 6.

Figure 6. The correlation between sample size and effect size

It was observed that the slope of the line increased towards the positive direction with the increase in the sample size. The statistical results regarding the slope are given in Table 7.

Table 7. Statistical results for the sample size and the effect size

<table>
<thead>
<tr>
<th></th>
<th>Point Estimation</th>
<th>Standard Error</th>
<th>Lower Limit</th>
<th>Upper Limit</th>
<th>Z value</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slope</td>
<td>0.00025</td>
<td>0.00003</td>
<td>0.00020</td>
<td>0.00030</td>
<td>9.13492</td>
<td>0.00000*</td>
</tr>
<tr>
<td>Intercept</td>
<td>-0.69905</td>
<td>0.02476</td>
<td>-0.74758</td>
<td>-0.65052</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

According to the results, it was observed that the increase in the sample size caused an increase of .025 in the effect size and this increase was statistically significant (p<.05).
Discussion

In the last decade, many researchers examined the association between FLA and EFL performances. This particular study aimed to systematically synthesize the results of those studies through meta-analysis. The first research question was about the direction and the extent of the relationship between EFL students’ FLA and their language performance. In this context, the first critical step was to identify the direction of the association between FLA and language performance. The results revealed a negative direction, which is parallel to the findings of the other studies in the literature (Horwitz, 2000; Gregersen, 2003; Batumlu & Erden, 2007; Tugan, 2015; Oruç, 2020). Al-Saraj (2014) posited that FLA negatively influences students’ language learning achievement or performance, which was supported with this study. The second critical step was to identify the correlation between the FLA and language performance. According to the results, the correlation between the two structures was found to be -.61. This value was between -.52 and -.71 with 95% confidence interval. Considering Thalheimer and Cook’s (2002) guidelines for interpreting effect sizes, it was considered as a medium effect. In addition, FLA explained 37% of the total variance of language performance. In order to better interpret the findings, the results were compared with the results of the other meta-analysis studies on the same subject. The effect size found in this study was higher than the ones in Horwitz’s study (2000) (r = −.50), Teimouri, Goetze and Plonsky’s study (2019) (r = −.47), and Zhang’s study (2019) (r = −.34). Therefore, it is concluded that FLA significantly affects EFL learners’ performances.

It is seen that the examined studies’ contribution levels and weighted effect sizes to the overall effect size were different (see the forest plow). Specifically, while some studies reported a negative effect of FLA on EFL performance (Serraj & Noordin, 2013; Hamzavi & Afšar 2014; Lu & Li, 2015; Özkan, 2019), there exist other studies reported an opposite result (Demirdaş & Bozdoğan, 2013; Lian & Budin, 2014; Tridinanti, 2018; Hadinata, 2020). In regard to these conflicting results, the second research question was to examine the effect of FLA on EFL performance in terms of some moderators.

The first moderator was grade level and the results revealed its non-significant moderator effect on the association between FLA and EFL performance. In a study, Tuncer and Akmençe (2018) compared high school students with students in higher education programs and found no difference between the groups in terms of school level, which supports the findings of the current study. Can and Can (2014) argued that the issues in learning foreign language are due to the combination of many structures related to students, teachers/instructors, curriculum, administrators, and parents. The fact that the grade level was not a significant moderator between FLA and EFL performance points out similar issues at different grade levels. In the present study, it was determined that FLA has a medium level effect on students’ EFL performance both in high school and higher education.

Due to the role of culture on the formation of anxiety (Yoğurtçu & Yoğurtçu, 2013), country in which studies were conducted was considered as a moderator variable in this particular study. The results revealed that country had significant moderator effect on the association between FLA and language performance. Cultural background influences individuals’ feelings, experiences, and reflections. Therefore, culture is considered as a powerful variable that affects anxiety level (Hofmann & Hinton, 2014). The significant moderator effect of country on the association between FLA and EFL performance in the present study may be due to cultural differences.

Another moderator variable of the current study was the type of anxiety (speaking, listening, writing, reading, and test anxiety). The findings proved a significant moderator effect of type of anxiety on the association between FLA on EFL performance. Specifically, it was observed that listening, test, speaking, reading, and writing anxiety affected EFL learning performance respectively from the most to the least. Zhang (2019) reported a similar finding and stated that this may be related to the nature of listening. Listening is a cognitive task that involves various complex structures of linguistic, semantic, and pragmatic knowledge (Rost, 2011) that consist of a series of sequential audio inputs (Vandergrift, 2011). In addition, listening is among the skills that challenge cognitive processes the most (Goh, 2000). Due to the complex structure of listening, it may cause learners to be affected by anxiety at the highest level.

Publication year was considered as a moderator variable in this study. According to the results, publication year was a predictor of the negative association between FLA and EFL performance. On the other hand, Zhang (2019) found that publication year did not predict this negative association and concluded that the degree of the correlation was consistent over the past decades. Considering that the current study included more recent studies, it is concluded that the negative association between FLA and EFL performances has increased in recent years.
The last moderator variable in the study was the sample size. According to the results, an increase in sample size resulted in a decrease in anxiety level. Sadık and Nasırcı (2019) posited that teachers and students may have higher anxiety level in less crowded classrooms since in those classrooms, teachers’ and students’ every movement are followed up by the others easily. This may cause teachers and students not to act in a natural way, and as a result, to increase their anxiety levels. Therefore, parallel to the findings in other studies in the literature, the finding of this study revealed that higher sample size decreases the level of anxiety and increases EFL performance.

Limitations and Directions for Future Research

This study has some limitations. According to the inclusion criteria of this particular study, only studies written in English or Turkish were selected for analysis, which may prevent to be reached out the actual effect size. Therefore, it is suggested future meta-analysis studies include studies written in other languages as well. The second limitation of the study is that a limited number of studies were reached at middle school level. Therefore, it is critical to include grade level as a moderator variable in future studies in order to determine the effects of FLA on EFL performance. A similar limitation exists for the country variable. In addition, the types of scales used in the studies included in the present study were not examined as moderators. Scale types used in studies can be important in explaining the relationship between FLA and EFL performance. Therefore, future research may examine the scale type as a moderator. There might be studies that were not published in the databases that eligible studies were searched in this particular study. Future studies should consider accessing different databases to find out more studies in this topic. In addition to those limitations, it is also critical to examine the association between FLA and EFL performance in terms of different moderator variables. Considering the negative effect of FLA on EFL performance, future studies should focus on the reasons of anxiety and intervention programs to develop students’ skills to manage anxiety in order to increase the effectiveness of language learning/teaching.

Conclusion

This meta-analysis study examined the association between FLA and EFL performance. Overall, the results revealed a medium level and negative effect between students’ FLA and EFL performance. Also, FLA negatively predicted EFL performance. In addition, the association between FLA and EFL performance was investigated based on some moderators. According to the results, FLA was found to cause similar issues in terms of EFL performance in high school and higher education levels. On the other hand, the association between FLA and EFL performance significantly differs in terms of the country moderator. Also, according to the results, the listening anxiety decreased EFL performance the most. In addition, sample size was found to be a significant moderator in the analysis. It is determined that crowded classrooms may decrease students’ FLA levels. As a result, the negative consequences of high FLA and low EFL performance for language learners are source of concern. Therefore, practitioners should create learning environments in which students feel comfortable and have less anxiety. It is also important to conduct extensive studies to identify pedagogical interventions that encourage students to alleviate their FLA. In addition, some measures can be taken to reduce listening anxiety in EFL learning in the classroom. Some of these measures can be applied in crowded classrooms, as demonstrated in the findings of the present study. Also, it is recommended to conduct studies patterned with qualitative research methods in order to examine the effects of culture on language learning.

Disclosure Statement

I wish to confirm that there are no known conflicts of interest associated with this publication and there has been no significant financial support for this work that could have influenced its outcome.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.
References

(* shows the studies included in the meta-analysis.)


Can, A. (2013). SPSS ile bilimsel araştırma sürecinde nicel veri analizi [Quantitative data analysis in the scientific research process with SPSS]. Ankara: Pegem Akademi Publisher


*Rezaabadi, O. T. (2016). The relationships between social class, listening test anxiety and test scores. Advances in Language and Literary Studies, 7(5), 147-156. https://doi.org/10.7575/aila.v.7n.5p.147


