# Relations Between Gender Stereotyping and Foreign Language Attainment: The Mediating Role of Language Learners' Anxiety and Self-Efficacy

Gulsah Kutuk[1], David William Putwain[1], Linda K. Kaye[2], Bethan Garrett[3]

[1]School of Education, Liverpool John Moores University, Liverpool, UK.

[2]Department of Psychology, Edge Hill University, Lancashire, UK

[3]Faculty of Education, Edge Hill University, Lancashire, UK

Data availability statement:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Acknowledgements:

This research was funded via a PhD studentship from Edge Hill University.

Address for correspondence: Gulsah Kutuk, School of Education, Liverpool John Moores University, IM Marsh Campus, Mossley Hill Rd, Liverpool, L17 6BD.

gulsah.kutuk@gmail.com

**Abstract**

**Background:** Gender stereotyping of academic domains has long been a major issue in education. However, previous research has mainly focused on male-dominated fields and women's disadvantage in such fields. Little attention has been paid to the fields of study, such as foreign language learning, which are typically stereotyped as female domains.

**Aims:** This study aimed to investigate whether relations between (1) learners' gender stereotypes about English as a foreign language (EFL) learning and language attainment, and (2) learner perceptions of teacher stereotypes of EFL learning and language attainment were mediated by anxiety and self-efficacy.

**Sample:** Data were collected from 701 university students ($M_{age}$ = 19.7 years, 49.4% male) learning EFL in three Turkish universities.

**Method:** Data were collected over three waves. Multi-group structural equation modelling approach was used to analyse the data.

**Results:** Results showed the relations between learners' gender stereotypes about EFL learning and language attainment was mediated by self-efficacy. Self-efficacy also mediated the relationship between learner perceptions of teacher stereotypes of EFL learning and language attainment, but only for women. Language anxiety was not a mediator between gender stereotypes and attainment in either model tested.

**Conclusions:** Findings show that gender stereotypes about EFL learning might affect learners' language attainment by altering their self-efficacy. Helping learners to maximise their self-efficacy will therefore be beneficial for their language attainment.

*Keywords:* Foreign language learning, gender stereotypes, self-efficacy, foreign language anxiety, language attainment

**Relations Between Gender Stereotyping and Foreign Language Attainment: The Mediating Role of Language Learners' Anxiety and Self-Efficacy**

From early adolescence, women and men start choosing different fields of study and occupations, often resulting in gender disparities in certain disciplines (Barone, 2011; van der Vleuten et al., 2016). While men are often interested in scientific or technical fields of education or careers, women are more likely to pursue careers that are more humanistic and care-oriented. Such imbalance could, in turn, be used to justify and reinforce gender stereotypes about academic ability (i.e., 'men are good at mathematics'; 'women are good at languages') that are not grounded in actual competencies. These stereotypes are also believed to increase with age (Hill & Lynch, 1983). Individuals who face increased pressure to conform to socially and culturally constructed gender roles experience an intensification of gender role through adolescence, particularly with respect to their achievement related beliefs (Usher & Pajares, 2008).

Women, for example, tend to report lower self-efficacy (i.e., perceived beliefs or judgments about one's own ability to succeed in a specific task or situation) than men, and believe that they do not have an innate ability in traditionally male-dominated fields such as Science, Technology, Engineering and Mathematics (STEM; Beyer, 2014; Tellhed et al., 2017). Likewise, it has been observed that women report higher mathematics anxiety compared to their male counterparts, and this leads to a corresponding decrease in their mathematics performance (e.g., Ganley & Vasilyeva, 2014; see Dowker et al., 2016). These findings suggest that individuals' achievement-related gender stereotypes may have a detrimental impact upon academic performance. The mechanisms underpinning this, however, require additional scrutiny and will be discussed in the following sections.

The extant research on the link between gender stereotypes and academic achievement has focused more or less exclusively on women and their disadvantage in male-dominated

disciplines. Relatively little research has been carried out on men and their performance in such areas as foreign language studies, even though this field has been often labelled as a female domain by students and teachers alike (see Carr & Pauwels, 2006; Chaffee, Lou, Noels, & Katz, 2020; Li et al., 2021; Pomerantz, 2008; Schmenk, 2004). In light of this, the overall aim of the current study was to examine the role of gender stereotypes in foreign language learning, a stereotypically female domain, drawing on Bandura's (1986) social cognitive theory (SCT). SCT construes human functioning as a triadic reciprocal interaction between personal (e.g., self-efficacy), environmental (e.g., social norms), and behavioural (e.g., achievement) processes. Accordingly, we first sought to explore the link between gender stereotypes of Turkish university students learning English as a foreign language (EFL) and their language attainment, and whether this link was mediated by anxiety and self-efficacy. We focused on anxiety and self-efficacy as personal factors in line with SCT. We theorised these to be especially relevant in the current Turkish educational context as the stakes are high for university-level EFL learners (i.e., they are under pressure to pass their qualifying examinations). Under these circumstances, learners' anxiety and self-efficacy play a crucial role in determining their success and failure (see Richardson et al., 2012; Üner et al., 2020).

Moreover, within SCT, socialisers' (i.e., significant others) gender stereotypes and attitudes are posited to have a significant impact on learners' achievement-related perceptions (Bandura, 1997; Bussey & Bandura, 1999). It is suggested that teachers'[1] expectations may influence students' ability-related perceptions and their actual performance in certain academic subjects (see Gunderson et al., 2012). Specifically, it is likely that teachers' implicit or explicit gendered stereotypes and expectations may lead learners to conform with these stereotypes and expectations. This potential effect is often labelled as 'self-fulfilling prophecy' (also known as Pygmalion effect) (Brophy, 1983; Cooper & Good, 1983;

Rosenthal & Jacobson, 1968). Given the age of the target population, we hypothesised that EFL learners in this study would be aware of their teachers' gender stereotypes and attitudes, and this might influence their achievement-related perceptions and subsequently their actual achievement. Therefore, the second aim of this study was to examine the relations between EFL teachers' gender stereotypes, which were assessed through learners' perceptions, and learners' language attainment, and whether anxiety and self-efficacy mediated this association.

**Gender Stereotypes and Their Association With Foreign Language Attainment**

Previous studies have demonstrated that foreign language learners generally perceive language learning as a female domain. For example, Carr and Pauwels (2006) interviewed over 200 boys (aged 12 to 18 years) studying a foreign language (e.g., French) in the major English-dominant communities such as Australia. Results showed that most participants believed language learning was not something 'real boys' do or are good at, and therefore avoided studying languages beyond compulsory level. Similarly, in a recent study, Chaffee, Lou, Noels, and Katz (2020) surveyed 1672 Canadian undergraduate students and found a stereotype of foreign language learning as feminine across different cultures and genders. Overall, these findings, along with previous evidence showing that men are consistently outperformed by women in language arts and foreign languages (Główka, 2014; Voyer & Voyer, 2014), suggest that men might be at a disadvantage in their language learning compared to women.

Nevertheless, it is important to note that gender, and how it is perceived by people, can change across culture, time, and different communities of practice (Eckert & McConnell-Ginet, 2003; Schmenk, 2004). This highlights the need for more research exploring the issue of gender stereotyping of foreign language learning with different groups of learners and identifying potential obstacles that might interfere with their achievement. The present study

is therefore timely since it aims to provide a fuller understanding of the phenomenon of gender stereotyping of foreign language learning by examining it in Turkey. Turkey is a non-English speaking country where feminine and masculine identities are well-defined and constructed in line with traditional gender roles and expectations (Bolak-Boratav et al., 2014).

A body of research shows that not only students, but also practicing and pre-service teachers hold gender stereotypes with respect to foreign language learning (Altan, 2012; Carr & Pauwels, 2006; Li et al., 2021; Peacock, 2001). Li et al. (2020), for instance, recently found that Chinese EFL teachers perceived their female students as more motivated, successful, and talented language learners than their male students. From a social cognitive perspective, teachers are important socialisation agents; their beliefs and attitudes can influence learners' own achievement-related beliefs and behaviours, and ultimately their achievement (Bandura, 1997). The plausibility of this assumption has already been investigated in some other domains such as mathematics (e.g., Lazarides & Watt, 2015). Despite the prevalence of teachers' gender stereotypes regarding foreign language learning, however, it is yet to be known how these stereotypes influence learners' educational outcomes in this particular domain. We, therefore, aimed to expand the existing research base by also investigating whether EFL learners' perceptions of their teachers' gender stereotypes predicted their educational outcomes above and beyond their own gender stereotypes.

Socialisers' influences can be measured from different perspectives (e.g., teachers' self-reports or students' reports). Many studies conducted to date has examined the relations between teachers' self-reported gender stereotypes and students' educational outcomes (e.g., Muntoni & Retelsdorf, 2018; Retelsdorf et al., 2015; Tiedemann, 2002). In the present study, however, we focused on students' reports to capture their perceptions of teacher stereotypes for two reasons. First, a core assumption of social cognitive theory is that socialisers' beliefs

and behaviours do not affect students' achievement-related outcomes directly but rather indirectly through students' subjective perceptions and interpretations of these beliefs and behaviours (Bandura, 1977). In other words, learners' perceptions and interpretations "form the bottle-neck of socializers' influence" (Gniewosz & Watt, 2017, p.1372). As such, our focus was not on teachers' gender stereotypes per se, but rather on what learners perceived their teachers' gender stereotypes to be. Second, teacher reports may be susceptible to social desirability and positive (or negative) self-presentation bias (Greenwald et al., 2009). Student reports may also be influenced by individual idiosyncrasies and other factors such as teacher popularity (De Jong & Westerhof, 2001). However, it is suggested that with a sufficient number of students in any one class, as is the case in the present study, student reports potentially offer a more accurate account than that of a single report from a teacher (Marsh & Roche, 1997).

**Mediating Roles of Self-Efficacy and Anxiety**

Evidence of the relationship between anxiety, self-efficacy, and attainment is widespread and well-documented in the field of second language acquisition. Language anxiety is defined as the worry and negative emotional reactions which learners experience when learning or using a second language (MacIntyre, 1999). As one of the most frequently experienced emotions in foreign language contexts, anxiety has been consistently found to be debilitating for different groups of learners in a wide range of countries (see Teimouri et al., 2019; Zhang, 2019, for two recent meta-analyses). Self-efficacy refers to perceived beliefs or judgments about one's own ability to succeed in a specific task or situation, and it can control human functioning through cognitive, motivational, affective, and decisional processes (Bandura, 1997). Strong positive correlations between self-efficacy and language proficiency have been observed in several investigations (e.g., Anam & Stracke, 2020). Some researchers have also examined the extent to which anxiety predicts language achievement alongside

self-efficacy and shown that there is a significant negative relationship between anxiety and self-efficacy (MacIntyre et al., 1997; Mills et al., 2006) suggesting that it is worthwhile to study these two constructs in conjunction with each other.

Prior evidence has shown that there are gender differences in self-efficacy and anxiety in domains that are gender stereotypical (Beyer, 2014; Ganley & Vasilyeva, 2014; Tellhed et al., 2017). In a meta-analysis, Huang (2013) has demonstrated that within different age groups (i.e., from 6 to over 23 years), men had higher self-efficacy in mathematics and sciences, while women had higher self-efficacy in the domain of language arts; the largest effect size was observed for respondents aged over 23 years old. Findings relating to the role of gender in foreign language anxiety and self-efficacy, however, are mixed and do not always align with the stereotypical expectations that men are less self-efficacious and more anxious in language learning compared to women (e.g., Dewaele et al., 2016; Mills et al., 2006; Park & French, 2013). Dewaele et al. (2016), for example, surveyed 1736 foreign language learners (aged 11-75 years) to determine the extent to which women and men learners differed from each other in terms of their enjoyment and anxiety. The results showed that although women reported that they enjoyed language learning more than men, they still experienced more anxiety compared to their male counterparts. In another study, Mills et al. (2006) showed that French listening self-efficacy related positively to listening proficiency only for undergraduate women, but not for men.

These findings support the theoretical assumption that self-efficacy is not shaped by gender per se but through the gender socialisation processes (Mills, 2014; also see Pajares & Usher, 2008), and therefore, learners' gender stereotypes, rather than their gender, need to be considered when exploring learners' achievement and achievement-related perceptions. Such empirical evidence is currently non-existent in the EFL learning literature. Therefore, using Bandura's (1986) social cognitive theory as a framework, we aim to examine the extent to

which learners' gender stereotypes and learner perceptions of teacher stereotypes relate to learners' self-efficacy and anxiety and subsequently influence their language attainment.

## The Present Study

The present study was guided by the research questions and hypotheses presented below:

1. *Research Question 1*: Do Turkish EFL learners' gender stereotypes about EFL learning relate to their language attainment through self-efficacy and anxiety?

*Hypothesis 1a*: A stereotype that EFL learning is a female domain is related to lower achievement in men, mediated by greater anxiety.

*Hypothesis 1b*: A stereotype that EFL learning is a female domain is related to lower achievement in men, mediated by lower self-efficacy.

*Hypothesis 2a*: A stereotype that EFL learning is a female domain is related to higher achievement in women, mediated by lower anxiety.

*Hypothesis 2b*: A stereotype that EFL learning is a female domain is related to higher achievement in women, mediated by higher self-efficacy.

2. *Research Question 2*: Do Turkish EFL learners' perceptions of their teachers' gender stereotypes about EFL learning relate to their language attainment through self-efficacy and anxiety?

*Hypothesis 3a*: A perception that language teachers believe EFL learning is a female domain is related to lower achievement in men, mediated by greater anxiety.

*Hypothesis 3b*: A perception that language teachers believe EFL learning is a female domain is related to lower achievement in men, mediated by lower self-efficacy.

*Hypothesis 4a*: A perception that language teachers believe EFL learning is a female domain is related to higher achievement in women, mediated by lower anxiety.

*Hypothesis 4b*: A perception that language teachers believe EFL learning is a female domain is related to higher achievement in women, mediated by higher self-efficacy.

## Method

### Participants

A total of 701 Turkish EFL learners were recruited using convenience sampling from two state (44.4%) and one private (56.4%) universities in Istanbul, Turkey. There was a fairly even split between men and women in the sample (men = 49.4%; women = 50.6%) with a mean age of 19.17 years (*SD* = 1.9). Participants were drawn from diverse backgrounds representing the seven regions of Turkey: Black Sea Region = 57.5%; Marmara Region = 7.3%; Aegean Region = 5.8%; Mediterranean Region = 6.3%; Eastern Anatolia Region = 3.9%; Central Anatolia Region = 3.4% and Southeastern Region = 1.4%. Participants were attending a one-year English preparatory course as a prerequisite for those aiming to pursue majors implementing the use of English as a medium of instruction. The EFL instruction in these courses was designed using the Common European Framework of Reference for Languages: Learning, Teaching, Assessment (CEFR) (see West et al., 2015 for further details). Aiming to standardise language teaching, learning and assessment across Europe, the CEFR describes language ability on a scale of six ascending levels (A1 – C2). At the time of data collection, 80.8% of the participants had been studying EFL for at least 6 months at university and they reported their levels as A2 – Elementary (2.8%), B1 – Pre-Intermediate (37.7%), B2 – Intermediate (52.0%), and B2 – Upper Intermediate (7.6%).

The data were collected over three waves using a prospective design (see Procedure for details). In the current data set, sixteen percent of data were missing due to participants being absent from their class or deciding not to participate at some point over the three waves of data collection. The randomness of missing data was tested using Little's missing completely at random (MCAR) test. Since Little's test was statistically significant ($p < .001$), MCAR could not be assumed. Further analyses revealed that missingness was due to a higher proportion of men with greater anxiety and lower self-efficacy being absent in the final time

point (i.e., EFL attainment scores). When the variables causing missingness (i.e., anxiety and self-efficacy) are included in all analytic models, data can be treated as missing at random (MAR) and handled using full information maximum likelihood (FIML) in M*plus* (Nicholson et al., 2017). FIML is effective in producing unbiased estimates since it addresses the issues arising from the missing data in longitudinal studies (Jeličić et al., 2009) even when the number of missing values is high (Enders, 2010).

## Measures

### *Questionnaire of Gender Stereotypes in Language Learning (QGSL)*

Gender stereotypes were assessed using a two-part questionnaire. The first part concentrated exclusively on participants' gender stereotypes about learning EFL. Participants were provided with five items such as "Please indicate which gender is generally good at learning English". The second part was concerned with learner perceptions of teacher stereotypes. The items used to assess learners' gender stereotypes in the first section were adapted to address learner perceptions of teacher stereotypes. For example, the above item was adapted as "Please indicate which gender is generally good at learning English according to your English teachers" (see Supplementary Materials for full list of items). Participants responded to items on a five-point scale (1 = *Always women*, 5 = *Always men*) with a neutral value of 3 (i.e., *Both men and women*). A mean score was computed with a low score indicating high female association. Internal consistency was good (McDonald's $\omega$ = .76 and .79, respectively for learners' stereotypes and learner perceptions of teacher stereotypes).

### *Multidimensional Language Class Anxiety Scale (MLCAS: Kutuk et al., 2020)*

The MLCAS is a 30-item scale intended to provide a multidimensional assessment of foreign language anxiety (Kutuk et al., 2020). It consists of a single common factor that represents EFL class anxiety as well as the major domains of language class anxiety (i.e., listening, speaking, reading, writing and testing). It also accounts for the variance due to the

cognitive, affective, and physiological components of anxiety. A sample item for affective listening activity anxiety is "Thinking about listening activities in class makes me feel uneasy" (see Supplementary Materials for full list of items). Participants responded to items on a 5-point scale (1 = *Strongly Disagree*, 5 = *Strongly Agree*). A mean score was computed with a high score representing a high level of anxiety. In the present study, internal consistency was excellent for the overall scale (McDonald's ω = .96).

***Questionnaire of Self-efficacy in Learning a Foreign Language (QSLL: [BLINDED])***

The QSLL is an 11-item questionnaire which was designed to assess participants' EFL self-efficacy. It consists of a single common factor that represents overall EFL self-efficacy as well as two specific group factors, namely receptive (i.e., reading and listening) and productive (i.e., speaking and writing) language skills. Questionnaire items were designed based on the Common European Framework of Reference for Languages: Learning, Teaching, Assessment which is an international standard for describing language ability (Council of Europe, 2001). A sample listening item is "I can watch and understand English films and TV series without English/Turkish subtitles" (see Supplementary Materials for the full list of the items). The items were scored on a 5-point scale (1 = *Strongly Disagree*, 5 = *Strongly Agree*). A mean score was computed, with a high score representing high self-efficacy. Internal consistency was excellent (McDonald's ω = .88).

***Confirmatory Factor Analysis***

In order to confirm the underlying factor structure of each of the three scales, a confirmatory factor analysis (CFA) was performed (see Table 1 for model fit indices and factor loadings). The CFA results supported the hypothesised one-factor structure of each part of the QGSL (i.e., Part 1, learner stereotypes; Part 2, learner perceptions of teacher stereotypes). Regarding the MLCAS and the QSLL, the results of the CFA revealed adequate fit to the data for bifactor models. A bifactor approach makes it possible to identify a single

general factor together with a number of specific orthogonal (i.e., uncorrelated) group factors (Reise et al., 2010). In the MLCAS, we used the EFL class anxiety as the general factor, and anxiety relating to listening, speaking, reading, writing and testing as the group factors (see Kutuk et al., 2020 for further details). In the QSLL, EFL self-efficacy was the general factor and self-efficacy relating to receptive and productive language skills were the two group factors. EFL class anxiety and EFL self-efficacy (i.e., the general factors) were of primary importance as evidenced by the items in the two scales displaying higher factor loadings on these factors than on their respective group factors (Hyland, 2015; Reise et al., 2010). They were, therefore, modelled as the overarching constructs (i.e., correlated parallel mediators) in the present study (see Figure 1).

<Insert Table 1 here>

***Foreign Language Attainment***

Learners' attainment was assessed using participants' average English examination scores which they received at the end of the English preparatory programme. The scores were calculated by the universities themselves based on a number of quizzes, mid-term and end-of-year exams. The structure of the exams they used to assess the English proficiency was similar across the universities by virtue of using the CEFR as a common assessment framework. Participants were assessed for their reading, writing, listening and speaking competencies which constituted one final score. The maximum score that the participants could get was 100%. The examinations were prepared by an independent testing office in each university. The testing offices comprised experienced EFL teachers who were responsible for the content, preparation and implementation of the examinations to be held throughout the academic year. Examinations were double-marked internally using the guidelines provided by the testing offices. Any discrepancies between the grades given by

two independent markers were discussed between the markers and a moderator, and a final single grade was determined with the agreement of all parties.

**Procedure**

Data collection was conducted in three phases between March and July in 2017. First, participants were asked to complete a demographics questionnaire and the QGSL in paper-and-pencil form. It was believed that the QGSL might create a priming effect and unconsciously limit the scope of participants' responses in subsequent questionnaires. To avoid this, the second phase began approximately 4 weeks following the end of the first phase. In the second phase, participants completed the MLCAS and the QSLL. The third and last phase was the collection of participants' language attainment scores which lasted from June to July 2017. Permission to conduct the study was first gained from the directors of the School of Foreign Languages offering the English preparatory courses and the teachers willing to participate in our study in these courses. Participants were also asked to give their individual consent on the first page of the demographic questionnaire. The project was approved by a Faculty Research Ethics Committee.

**Analytic Strategy**

Except where stated, analyses were conducted in M*plus* 8.3 using the maximum likelihood estimator with robust standard errors (MLR) to adjust standard errors and chi-square goodness of fit statistics to account for the non-normal distribution of data (Muthén & Muthén, 2019). Data analysis was performed in three steps. First, we conducted a number of preliminary analyses to (1) screen the data for outliers and test the multivariate normality assumption of the data, (2) provide descriptive statistics, and (3) determine intercorrelations amongst the study variables. Second, a measurement model consisting of five latent factors (i.e., learners' stereotypes, learner perceptions of teacher stereotypes, anxiety, self-efficacy and attainment) was tested using a multiple group (i.e., women and men) analysis. We

employed a CFA to examine the properties of the measurement model, test measurement invariance, and to estimate latent bivariate correlations. Third, we used structural equation modelling (SEM) to examine the mediator effects of anxiety and self-efficacy on the relations between gender stereotypes and language attainment using a multiple group analysis (Figure 1). The analyses were run using the 'type = complex' command in M*plus* to adjust standard errors for the clustering of data within schools (Wu & Kwok, 2012).

<Insert Figure 1 here>

## Results

### *Preliminary Analyses*

The data were first screened for univariate and multivariate outliers. To reduce the influence of any outliers detected, data points which were 3 *SD* beyond the corresponding means were winsorized (replaced by a *Z* score of +3 or -3). The multivariate normality assumption was also investigated through Mardia's two-sided test of fit for skewness and kurtosis (Wang & Wang, 2019). Mardia's test was statistically significant ($p$ <.001) indicating that the data did not meet the criteria for multivariate normality. To deal with the non-normality, we used the MLR estimator in M*plus*.

Descriptive statistics and Pearson correlations between the study variables were calculated using SPSS v25 and presented in Table 2 and 3. The bivariate correlations revealed no issues of multicollinearity. Collinearity among the predictor variables was further examined using tolerance and the variance inflation factor (VIF) values which were all within acceptable limits (Tolerance > .67; VIF < 1.4).

<Insert Table 2 here>

<Insert Table 3 here>

### *Measurement Model*

Model fit was established using the Root Mean Square Error of Approximation (RMSEA), the Standardised Root Mean Square Residual (SRMR), Comparative Fit Index (CFI), and the Tucker-Lewis index (TLI). A good model is indicated by RMSEA ≈ .05, SRMR ≈ .08, and CFI and TLI ≈ .95 (Hu & Bentler, 1999; Marsh et al., 2004). By these criteria, our measurement model yielded a relatively good fit, $\chi^2(2112) = 348.504$, $p < .001$, RMSEA < .039, SRMR = .049, CFI = .937, and TLI = .924. The standardized estimates of factor loadings for the study constructs were within the acceptable range (.41–.79), suggesting good construct validity.

***Measurement Invariance***

Establishing measurement invariance is a prerequisite for comparing different groups (Chen, 2007). To ensure data from female and male participants were comparable, we tested measurement invariance with a series of multi-group CFA models in four steps: configural invariance, metric (or weak) invariance, scalar (or strong) invariance and residual (or strict) invariance. Configural invariance tests whether the factor structures of the measures are equivalent across groups. This is followed by the subsequent steps where factor loadings, item intercepts, and item residuals are constrained to be equal across groups respectively for metric invariance, scalar invariance and residual invariance. Invariance is supported if changes in model fit statistics are within recommended cut-off values (i.e., ΔRMSEA is <.015 and ΔCFI and ΔTLI are <.01) (Chen, 2007).

Results are presented in Table 4. Overall, our measures were invariant across gender. All of them achieved residual invariance except for the QGSL assessing learner perceptions of teacher stereotypes. It showed partial scalar invariance after freeing the constraint for the intercept on one item. Since partial scalar invariance is deemed sufficient for comparing relations between variables (Gregorich, 2006), it was decided that a sound psychometric basis for comparing the data from women and men was evident.

<Insert Table 4 here>

## Structural Equation Modelling

### *Learners' Gender Stereotypes*

The model provided a reasonable fit to the data: $\chi^2(1803) = 2905.935$, $p < .001$, RMSEA < .04, SRMR = .053, CFI = .934, and TLI = .921. The results for the path analyses are shown in Figure 2 and Table 5. Paths from learners' gender stereotypes to self-efficacy, and from self-efficacy to language attainment were statistically significant. Women who believed that EFL learning is a female domain reported higher self-efficacy, and men with the same stereotype reported lower self-efficacy. A greater self-efficacy was strongly associated with better language attainment for both women and men. Learners' gender stereotypes were not directly related to language attainment for either gender. Learners' gender stereotypes were unrelated to anxiety, and anxiety was unrelated to attainment.

The indirect paths from gender stereotypes to language attainment were assessed by creating 95% confidence intervals (CIs) around a point estimate. When 95% CIs do not cross zero, indirect paths are statistically significant at $p$ <.05 (MacKinnon, 2012). The total indirect paths from women's and men's gender stereotypes to language attainment were statistically significant. The specific indirect paths showed that the relationship between learners' gender stereotypes and language attainment was mediated by self-efficacy. More specifically, women who held the stereotype that EFL learning is a female domain had higher attainment mediated by higher self-efficacy. In contrast, men with the same stereotype had lower attainment mediated by lower self-efficacy. Indirect paths from learners' gender stereotypes to attainment, mediated by anxiety, were not statistically significant.

<Insert Figure 2 here >

### *Learner Perceptions of Teacher Stereotypes*

The model yielded an acceptable fit to the data: $\chi^2(1801) = 3022.9868$, $p < 001$, RMSEA < .044, SRMR = .054, CFI = .929, and TLI = .914. The results for the path analyses are shown in Figure 3 and Table 5. Paths from learner perceptions of teacher stereotypes to anxiety and self-efficacy were statistically significant, but only for women. Women's perceived teacher stereotypes were negatively associated with anxiety and positively with self-efficacy. Greater self-efficacy was associated with higher language attainment for both genders. Learner perceptions of teacher stereotypes were not directly related to language attainment for either women or men. Anxiety was unrelated to attainment.

The total indirect path from learner perceptions of teacher stereotypes to language attainment was statistically significant only for women. The specific indirect paths showed that the relation between learner perceptions of teacher stereotypes and their attainment was mediated by self-efficacy, but not anxiety. In other words, women who perceived their teachers to hold the stereotype that EFL learning is a female domain performed better due to higher self-efficacy.

<Insert Figure 3 here>

<Insert Table 5 here >

## Discussion

The present research investigated whether Turkish university students' gender stereotypes (i.e., learner stereotypes and learner perceptions of teacher stereotypes) in EFL learning were associated with their language attainment, and whether this association was mediated by self-efficacy and anxiety. The study findings supported *Hypothesis 1b* and *Hypothesis 2b* in that the relations between learners' gender stereotypes and language attainment were mediated by self-efficacy both for women and men. That is, self-efficacy served as the mechanism through which learners' gender stereotypes showed positive or negative relations with language attainment. From a theoretical perspective, this is consistent

with SCT which proposes that how people acquire and maintain certain behavioural patterns is determined by a triadic interaction between personal, environmental, and behavioural processes (Bandura, 1997).

More specifically, our findings suggest that when women believe EFL learning is a female domain, their judgements about their abilities to accomplish the necessary language tasks are more likely to be positive. Such positive judgements, in turn, help women to achieve better learning outcomes. This may be because these women, as individuals with stronger self-efficacy, would tend to invest more time and effort in language learning and be more determined to overcome the challenges and setbacks that inevitably arise (Bandura, 1994). On the other hand, the same stereotype (i.e., that EFL learning is a female domain) lowers the strength of men's belief that they can achieve desired outcomes in language learning. A male learner with low self-efficacy may therefore invest less time and effort in accomplishing a language task, demonstrate less perseverance when faced with challenges or perceive some tasks to be more difficult than they actually are (see Schunk & Pajares, 2009). Less effort and persistence would, in turn, negatively affect men's achievement outcomes in EFL. In a nutshell, the more efficacious students feel about their learning due to their gender stereotypes, the more likely they are to engage with and attain their language learning and development goals. With these findings, our study expands upon previous research (Carr & Pauwels, 2006; Chaffee, Lou, & Noels, 2020; Chaffee, Lou, Noels, & Katz, 2020; Pomerantz, 2008; Schmenk, 2004) in showing how gender stereotypes are relevant for EFL learners' self-efficacy and anxiety, and therefore their actual language attainment.

*Hypothesis 3b* was also supported by the study findings which revealed that women who perceived that their teachers believed EFL learning is a female domain had an increased self-efficacy and therefore obtained a higher attainment score. The present finding is in line with the theoretical assumption that socialisers' gender stereotypes and attitudes play a

crucial role in the development of learners' achievement-related perceptions (Bandura, 1997; Bussey & Bandura, 1999). It is also consistent with a more general literature demonstrating that teachers' gender stereotypes might predict girls' achievement in stereotypically female domains (Muntoni & Retelsdorf, 2018). Although further work is needed to fully understand the implications of our study findings, it is possible that a positive self-fulfilling prophecy may be operating in foreign language learning contexts. That is, teachers' stereotypes favouring women in EFL learning may lead women to act in accordance with these stereotypes (e.g., become more motivated to learn EFL), eventually making the stereotypes come true.

Mediation of the relationship between learner perceptions of teacher stereotypes and language attainment by self-efficacy suggests that teachers' gender stereotypes do not influence language attainment directly. Rather, they alter women's beliefs about their capabilities to perform language learning tasks successfully at designated levels. Addressing the recent call for research exploring how learners' learning environment and achievement-related perceptions might be shaped by language teachers' gender stereotypes (Li et al., 2021), our study demonstrates that self-efficacy is at least one important mechanism by which teachers' stereotypes might be related to women's foreign language attainment. It is possible that in contexts where EFL learning is perceived to be a female domain by teachers by teachers, women are in a more advantageous position than men in terms of forming and developing stronger self-efficacy.

This may be because in these contexts, there are more opportunities for women to construct their self-efficacy from the main sources of information (i.e., enactive mastery experiences, vicarious experiences, verbal persuasion, and physiological and affective states). For example, in a classroom environment, women might receive consistent positive feedback (e.g., "I know you can do it") from their teachers highlighting their capabilities and

performance. Even though such verbal persuasion does not necessarily create a long-lasting effect on self-efficacy, it has the potential to help women to make an increased effort or try hard enough to succeed (Bandura, 1997). Also, the impact of the persuasion can vary depending on the credibility, trustworthiness, and expertise of the persuader. Given that in most foreign language contexts, teachers are often regarded as the main source of information, women are likely to trust their teachers more than other people. Thus, any positive verbal encouragement received from them would further enhance their competence beliefs, resulting in better learning outcomes. To a certain extent, this may account for the female 'superiority' in foreign language learning that has been reported in the Turkish context (Dayioğlu & Türüt-Aşik, 2007) and elsewhere (Główka, 2014; Voyer & Voyer, 2014).

Contrary to our expectation, however, *Hypothesis 4b* was not supported by our study findings. Learner perceptions of teacher stereotypes did not predict men's language attainment through a weakened self-efficacy although there was a statistically significant link between self-efficacy and attainment. These findings suggest that the self-fulfilling prophecy may not exist for male EFL learners in the current context. This is somewhat inconsistent with some previous work (Retelsdorf et al., 2015; Wolter et al., 2015) which found that teachers' gender stereotypes predicted boys' lower self-concept, motivation, and achievement in stereotypically female domains. Nevertheless, contradictory findings offer support for the ideas of Brophy (1983) who suggested that self-fulfilling prophecies do not automatically occur. Rather, there are a number of factors such as socio-economic status, ethnicity, age, and motivation influencing learners' susceptibility to teachers' gender stereotypes (De Boer et al., 2010).

To explain our unexpected results, we put forward two considerations. One possibility is that in the current study context, differences in socialisation processes played a role in the

formation of learners' self-beliefs which mirrors Bandura's (1997) theorising that cultural values and practices are likely to determine how efficacy beliefs are developed and related to performance. In Turkey, men are generally associated with authority and power while women are expected to be loyal and obedient and seen as in need of protection and care (Boratav et al., 2014). Given their status in society, it is possible that Turkish men are less susceptible to others' negative gender stereotypes about themselves as such and, therefore, their subject-related efficacy and attainment are not as strongly influenced by these stereotypes compared to their female peers. This is in line with SCT which asserts that learners are not passively shaped by their learning environment (Usher & Schunk, 2018). Rather, by exerting their human agency, learners can adapt and manage their learning environment.

A second possibility is that other factors (e.g., values attributed to EFL learning by men) played a moderating role in the link between gender stereotypes and self-efficacy. SCT contends that perceived importance and utility of learning are important determinants of individuals' actions (Bandura, 1997; Schunk & DiBenedetto, 2020; Schunk & Usher, 2019). For the participants in the present study, it was compulsory to attend the English classes and pass the final English proficiency test to be able to start studying their chosen academic major. Poor English performance would mean losing both time and money for these learners and their families since they would have to study English for another year if they failed. In light of this, we suspect that men's perceived value of EFL learning could have shielded their self-efficacy against the destructive effects of teachers' gender stereotypes. For women, on the other hand, high perceived value may have increased the strength of the relationship between teachers' stereotypes favouring women in EFL and women's self-efficacy which was already significant.

Evidence was not obtained in support of our hypotheses regarding the mediating role of anxiety (*Hypotheses 1a, 2a, 3a, and 4a*). The only significant link was observed between

learner perceptions of teacher stereotypes and anxiety among women. Women who perceived that their teachers believed EFL learning is a female domain had lower anxiety. However, this decrease did not lead to an increase in their attainment. The results showed that anxiety was highly negatively correlated with self-efficacy for women and men, which mirrors those of the previous studies (MacIntyre et al., 1997; Mills et al., 2006). Consistent with SCT (Bandura, 1986, 1997), this suggests that learners with higher self-efficacy were less anxious. Given the strong relation between anxiety and self-efficacy, we suspect that there was insufficient unique variance for anxiety to predict the participants' attainment. Thus, when the shared variance between anxiety and self-efficacy is considered, self-efficacy emerges as the mechanism through which the gender stereotypes in question exert their effects on learners' language attainment.

Our study makes an important contribution because it is the first to investigate the aforementioned relations in Turkey where traditional gender roles and expectations prevail. Although it is not possible to ascertain the extent to which findings are transferable to other international contexts, we anticipate that they can be generalised to populations with similar cultural and social backgrounds. Theoretically, it would be useful to explore the phenomenon of gender stereotyping for foreign language learning in more egalitarian societies. We will therefore develop a better understanding of the scope and nature of gender stereotypes in relation to foreign language learning in different contexts and among learners with different cultural and social backgrounds.

**Limitations and Directions for Future Research**

The results of the present study need to be considered in light of the following limitations. First, self-efficacy and anxiety were measured at one time point only due to the limited time frame and the difficulty accessing the same participants over a period of time. Given that self-beliefs are dynamic (Bandura, 1986; Dörnyei, 2000), the mediating role of such constructs

could be investigated in future studies by adopting a longitudinal design in which their changing levels are examined over time. Second, due to the correlational nature of the results no causation can be inferred. Researchers seeking a causal explanation should adopt an experimental approach such as using stereotype threat manipulations (e.g., Chaffee, Lou, & Noels, 2020; Keller, 2007) and investigate the effects of gender stereotypes on EFL achievement accordingly.

Third, the current study was concerned with learner perceptions of teacher stereotypes rather than language teachers' actual gender stereotypes. As Gniewosz and Watt (2017) asserted, there are conceptual differences between perceived and actual representations of reality, and they should be treated as such. In line with this assertion, numerous studies have documented that there are considerable discrepancies between student and teacher ratings in relation to, for example, instructional quality (Kunter & Baumert, 2006; Wagner et al., 2016) and classroom environment (Spearman & Watt, 2013). It is therefore suggested that in future studies, researchers choose their data source (e.g., whether teacher reports, student reports, or both) carefully considering their theoretical and methodological approaches. On a related note, through self-report questionnaires, we focused on learners' explicit (rather than implicit) gender stereotypes relating to foreign language learning. It is likely that in topics such as stereotyping, implicit attitudes and stereotypes might predict behaviour better than the explicit ones (see Greenwald et al., 2009). Therefore, future investigations should also explore learners' as well as teachers' implicit attitudes and stereotypes and thereby determine the mechanisms driving the relationship between gender stereotypes and language achievement.

Fourth, in the present study, foreign language attainment was measured using participants' average English examination scores since they had high ecological validity and could be obtained with minimal intrusiveness (Brown et al., 2018). However, we

acknowledge that using these scores as a measure of language attainment presents certain disadvantages (e.g., assessment literacy among language teachers may vary across universities). In future studies, such scores should be used in conjunction with other measures (e.g., standardised language assessments) to improve the validity of research findings (see Brown et al., 2018, for relevant discussion).

Finally, prior levels of English language proficiency were not controlled for in the present study. To rule out the potential effects of this variable, future studies should include prior levels of language proficiency in their analyses, and test whether they influence the nature of the relationship between gender stereotypes and language attainment via self-efficacy and anxiety. Language proficiency can be measured by formal measures such as a standardised language test or through informal methods such as teacher ratings.

**Practical Implications**

Notwithstanding these limitations, the current study has some important pedagogical implications. We found evidence that the stereotype that EFL learning is a female domain related to women's and men's language attainment through a heightened and decreased self-efficacy respectively. We also found that learner perceptions of teacher stereotypes had a positive effect on women's language attainment by boosting their self-efficacy. These findings indicate that when gender stereotypes favouring women in foreign languages are concerned, men are the disadvantaged group. Accordingly, we suggest that teachers adopt strategies to minimise the negative effects of such stereotypes on men's self-efficacy and language attainment. This can be achieved by first (a) raising teachers' awareness of gender stereotypes and their impact, and (b) providing teachers with a more detailed knowledge of the sources of efficacy and how to apply this knowledge in language classrooms effectively.

Consciously or not, teachers themselves may believe in gender stereotypes in language learning (e.g., Carr & Pauwels, 2006; Li et al., 2020) and therefore behave accordingly (e.g.,

grouping men and women for certain language tasks and activities) in language classrooms. These might activate learners' existing gender stereotypes subtly (or otherwise) or develop their knowledge of such stereotypes. Previous research suggests that, for example, when individuals perceive that their teachers have gender stereotypes, it is more likely that they endorse those stereotypes themselves (e.g., Keller, 2001). The internalised gender stereotypes might then impair men's achievement-related behaviours and outcomes. We therefore suggest that during their teacher education and continuing professional development, teachers are made aware of widely shared gender stereotypes and how these stereotypes might be relevant for their teaching practices and their students' achievement-related behaviours and outcomes. Having a deeper understanding of the nature of gender stereotypes and their potential impact on learning will inform teachers about how to alter their teaching practices and increase learning opportunities for the disadvantaged group, men.

Also, teachers should provide equitable opportunities for both men and women with diverse needs to form and develop their self-efficacy from various sources of information. They should pay special attention to, among other things, using verbal persuasion (i.e., verbal feedback and encouragement or discouragement) methods appropriately in language classrooms. If teachers, for example, provide more encouragement (e.g., praising women more than men) or set higher expectations of the positively stereotyped group (i.e., women), men might feel ignored or excluded in language learning environments. On the other hand, if teachers aim at increasing men's self-efficacy believing they are at a disadvantage, they may then leave out women who do not perceive themselves as capable of accomplishment. This may, in turn, reflect on learners' achievement-related behaviours and outcomes. As such, rather than aiming at enhancing a particular gender group's self-efficacy, teachers should monitor their individual students' progress closely and support those whose self-efficacy

needs to be enhanced by adopting relevant approaches (see Mills, 2014, for further information).

## Conclusion

The present study is the first to have prospectively examined relations between gender stereotypes in EFL learning and language attainment and provided a deeper understanding of the mechanisms at work in this relationship among university level EFL learners in a non-English speaking country. In general, our hypotheses regarding the mediating role of self-efficacy were supported by the data. Using a robust prospective design with temporal separation of measures, we found that when shared variance between anxiety and self-efficacy was considered, (1) learners' gender stereotypes were related to women's and men's attainment via their self-efficacy, (2) learner perceptions of teacher stereotypes were related to women's attainment through a heightened self-efficacy. At a more practical level, these results suggest that language teachers need to become aware of their students' gender stereotypes in EFL learning as well as their own potentially stereotyped beliefs and behaviours. They should then make a concerted effort to address any issues, such as decreased self-efficacy, that learners may experience due to these stereotypes.

## Endnotes

[1] In Turkey, English as a foreign language is taught by 'instructors' in a Higher Education context. However, for the sake of being consistent with the extant literature, we used the term 'teachers' instead of 'instructors' throughout the paper.

## References

Altan, M. Z. (2012). Pre-service EFL teachers' beliefs about foreign language learning. *European Journal of Teacher Education, 35*(4), 481–493. https://doi.org/10.1080/02619768.2011.643399

Anam, S., & Stracke, E. (2020). The role of self-efficacy beliefs in learning English as a foreign language among young Indonesians. *TESOL Journal, 11*(1), 1–21. https://doi.org/10.1002/tesj.440

Bandura, A. (1977). *Social learning theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.),Encyclopedia of human behavior (Vol. 4, pp. 71–81). Academic Press.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York, NY: Macmillan.

Barone, C. (2011). Some things never change: Gender segregation in higher education across eight nations and three decades. *Sociology of Education*, *84*(2), 157–176. https://doi.org/10.1177/0038040711402099

Beyer, S. (2014). Why are women underrepresented in Computer Science? Gender differences in stereotypes, self-efficacy, values, and interests and predictors of future CS course-taking and grades. *Computer Science Education*, *24*(2–3), 153–192. https://doi.org/10.1080/08993408.2014.963363

Bolak-Boratav, H., Fişek, G. O., & Eslen-Ziya, H. (2014). Unpacking masculinities in the context of social change: Internal complexities of the identities of married men in Turkey. *Men and Masculinities*, *17*(3), 299–324. https://doi.org/10.1177/1097184X14539511

Brophy, J. E. (1983). Research on the self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, *75*(5), 631–661. https://doi.org/10.1037/0022-0663.75.5.631

Brown, A. V., Plonsky, L., & Teimouri, Y. (2018). The use of course grades as metrics in L2 research: A systematic review. *Foreign Language Annals*, *51*(4), 763–778. https://doi.org/10.1111/flan.12370

Bussey, K., & Bandura, A. (1999). Social cognitive theory of gender development and differentiation. *Psychological Review*, *106*(4), 676-713. https://doi.org/10.1037/0033-295X.106.4.676

Carr, J., & Pauwels, A. (2006). *Boys and foreign language learning: Real boys don't do languages*. Palgrave Macmillan.

Chaffee, K. E., Lou, N. M., & Noels, K. A. (2020). Does stereotype threat affect men in language domains? *Frontiers in Psychology*, *11*, 1302. https://doi.org/10.3389/fpsyg.2020.01302

Chaffee, K. E., Lou, N. M., Noels, K. A., & Katz, J. W. (2020). Why don't "real men" learn languages? Masculinity threat and gender ideology suppress men's language learning motivation. *Group Processes & Intergroup Relations, 23*(2), 301–318. https://doi.org/10.1177/1368430219835025

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Cooper, H. M., & Good, T. L. (1983). *Pygmalion grows up: Studies in the expectation communication process*. New York, NY: Longman..

Council of Europe. (2001). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Cambridge University Press.

Dayioğlu, M., & Türüt-Aşik, S. (2007). Gender differences in academic performance in a large public university in Turkey. *Higher Education, 53*(2), 255–277. https://doi.org/10.1007/s10734-005-2464-6

De Boer, H., Bosker, R. J., & Van der Werf, M. P. (2010). Sustainability of teacher expectation bias effects on long-term student performance. *Journal of Educational Psychology, 102*(1), 168 -179. https://doi.org/10.1037/a0017289

De Jong, R., & Westerhof, K. J. (2001). The quality of student ratings of teacher behaviour. *Learning Environments Research*, *4*(1), 51–85.

Dewaele, J.-M., MacIntyre, P. D., Boudreau, C., & Dewaele, L. (2016). Do girls have all the fun? Anxiety and enjoyment in the foreign language classroom. *Theory and Practice of Second Language Acquisition, 2*(1), 41–63.

Dörnyei, Z. (2000). Motivation in action: Towards a process-oriented conceptualisation of student motivation. *British Journal of Educational Psychology*, *70*(4), 519–538. https://doi.org/10.1348/000709900158281

Dowker, A., Sarkar, A., & Looi, C. Y. (2016). Mathematics anxiety: What have we learned in 60 years? *Frontiers in Psychology, 7*, 508. https://doi.org/10.3389/fpsyg.2016.00508

Eckert, P., & McConnell-Ginet, S. (2003). *Language and gender*. NewYork: Cambridge University Press.

Enders, C. K. (2010). *Applied missing data analysis*. Guilford Press.

Ganley, C. M., & Vasilyeva, M. (2014). The role of anxiety and working memory in gender differences in mathematics. *Journal of Educational Psychology, 106*(1), 105-120. https://doi.org/10.1037/a0034099

Główka, D. (2014). The impact of gender on attainment in learning English as a foreign language. *Studies in Second Language Learning and Teaching*, *4*(4), 617–635. https://doi.org/10.14746/ssllt.2014.4.4.3

Gniewosz, B., & Watt, H. M. (2017). Adolescent-perceived parent and teacher overestimation of mathematics ability: Developmental implications for students' mathematics task values. *Developmental Psychology*, *53*(7), 1371–1383. https://doi.org/10.1037/dev0000332

Greenwald, A. G., Poehlman, T. A., Uhlmann, E. L., & Banaji, M. R. (2009). Understanding and using the Implicit Association Test: III. Meta-analysis of predictive validity. *Journal of Personality and Social Psychology, 97*(1), 17 –41. https://doi.org/10.1037/a0015575

Gregorich, S. E. (2006). Do self-report instruments allow meaningful comparisons across diverse population groups? Testing measurement invariance using the confirmatory factor analysis framework. *Medical Care*, *44*(11 Suppl 3), S78–S94. https://doi.org/10.1097/01.mlr.0000245454.12228.8f

Gunderson, E. A., Ramirez, G., Levine, S. C., & Beilock, S. L. (2012). The role of parents and teachers in the development of gender-related math attitudes. *Sex Roles, 66*(3–4), 153–166. https://doi.org/10.1007/s11199-011-9996-2

Hyland, P. (2015). Application of bifactor models in criminal psychology research: A guide to researchers. *Journal of Criminal Psychology*. 5(2), 65-74. https://doi.org/10.1108/JCP-03-2015-001

Hill, J. P., & Lynch, M. E. (1983). The intensification of gender-related role expectations during early adolescence. In J. Brooks-Gunn & A. C. Petersen (Eds.), *Girls at puberty* (pp. 201–228). Plenum.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Huang, C. (2013). Gender differences in academic self-efficacy: A meta-analysis. *European Journal of Psychology of Education*, *28*(1), 1–35. https://doi.org/10.1007/s10212-011-0097-y

Jeličić, H., Phelps, E., & Lerner, R. M. (2009). Use of missing data methods in longitudinal studies: The persistence of bad practices in developmental psychology. *Developmental Psychology*, *45*(4), 1195–1199. https://doi.org/10.1037/a0015665

Keller, C. (2001). Effect of teachers' stereotyping on students' stereotyping of mathematics as a male domain. *The Journal of Social Psychology*, *141*(2), 165–173. https://doi.org/10.1080/00224540109600544

Keller, J. (2007). Stereotype threat in classroom settings: The interactive effect of domain identification, task difficulty and stereotype threat on female students' maths

performance. *British Journal of Educational Psychology*, *77*(2), 323–338. https://doi.org/10.1348/000709906X113662

Kunter, M., & Baumert, J. (2006). Who is the expert? Construct and criteria validity of student and teacher ratings of instruction. *Learning Environments Research*, *9*(3), 231–251. https://doi.org/10.1007/s10984-006-9015-7

Kutuk, G., Putwain, D. W., Kaye, L., & Garrett, B. (2020). Development and validation of a new multidimensional language class anxiety scale. *Journal of Psychoeducational Assessment*, *38*(5), 649 –658. https://doi.org/10.1177/0734282919875881

Lazarides, R., & Watt, H. M. (2015). Girls' and boys' perceived mathematics teacher beliefs, classroom learning environments and mathematical career intentions. *Contemporary Educational Psychology*, *41*(2015), 51–61. https://doi.org/10.1016/j.cedpsych.2014.11.005

Li, J., McLellan, R., & Forbes, K. (2021). Investigating EFL teachers' gender-stereotypical beliefs about learners: A mixed-methods study. *Cambridge Journal of Education*, *51*(1), 19-44https://doi.org/10.1080/0305764X.2020.1772720

MacIntyre, P. D. (1999). Language anxiety: A review of the research for language teachers. In D. J. Young (Ed.), *Affect in foreign language and second language learning: A practical guide to creating a low-anxiety classroom atmosphere* (pp. 24–45). McGraw-Hill.

MacIntyre, P. D., Noels, K. A., & Clément, R. (1997). Biases in self-ratings of second language proficiency: The role of language anxiety. *Language Learning*, *47*(2), 265–287. https://doi.org/10.1111/0023-8333.81997008

MacKinnon, D. (2012). *Introduction to statistical mediation analysis*. Routledge.

Marsh, H. W., Hau, K.-T., & Wen, Z. (2004). In search of golden rules: Comment on hypothesis-testing approaches to setting cutoff values for fit indexes and dangers in

overgeneralizing Hu and Bentler's (1999) findings. *Structural Equation Modeling, 11*(3), 320–341. https://doi.org/10.1207/s15328007sem1103_2

Marsh, H. W., & Roche, L. A. (1997). Making students' evaluations of teaching effectiveness effective: The critical issues of validity, bias, and utility. *American Psychologist, 52*(11), 1187–1197. https://doi.org/10.1037/0003-066X.52.11.1187

Mills, N. (2014). Self-efficacy in second language acquisition. In S. Mercer & M. Williams (Eds.), *Multiple perspectives on the self in SLA* (pp. 6–22). Multilingual Matters.

Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self-efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals, 39*(2), 276–295. https://doi.org/10.1111/j.1944-9720.2006.tb02266.x

Muntoni, F., & Retelsdorf, J. (2018). Gender-specific teacher expectations in reading—The role of teachers' gender stereotypes. *Contemporary Educational Psychology, 54*, 212–220. https://doi.org/10.1016/j.cedpsych.2018.06.012

Muthén, L. K., & Muthén, B. O. (2019). Mplus Version 8 user's guide. *Muthén & Muthén*.

Nicholson, J. S., Deboeck, P. R., & Howard, W. (2017). Attrition in developmental psychology: A review of modern missing data reporting and practices. *International Journal of Behavioral Development, 41*(1), 143–153. https://doi.org/10.1177/0165025415618275

Pajares, F., & Usher, E. L. (2008). Self-efficacy, motivation, and achievement in school from the perspective of reciprocal determinism. In M. Maehr, T. C. Urdan, & S. Karabenick (Eds.), *Advances in motivation and achievement: Vol. 15. Social psychological perspectives* (pp. 391–423). Emerald Group Publishing.

Park, G.-P., & French, B. F. (2013). Gender differences in the foreign language classroom anxiety scale. *System, 41*(2), 462–471. https://doi.org/10.1016/j.system.2013.04.001

Peacock, M. (2001). Pre-service ESL teachers' beliefs about second language learning: A longitudinal study. *System, 29*(2), 177–195. https://doi.org/10.1016/S0346-251X(01)00010-0

Pomerantz, A. (2008). "Señora, boys just cannot speak Spanish": Negotiating masculine identities in a foreign language classroom. *Working Papers in Educational Linguistics (WPEL), 23*(2), 1-23. Retrieved from http://repository.upenn.edu/wpel/vol23/iss2/1

Reise, S. P., Moore, T. M., & Haviland, M. G. (2010). Bifactor models and rotations: Exploring the extent to which multidimensional data yield univocal scale scores. *Journal of Personality Assessment, 92*(6), 544–559. https://doi.org/10.1080/00223891.2010.496477

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology, 107*(1), 186-194. https://doi.org/10.1037/a0037107

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin, 138*(2), 353-387. https://doi.org/10.1037/a0026838

Rosenthal, R., & Jacobson, L. (1968). *Pygmalion in the classroom: Teacher expectation and pupils' intellectual development*. New York: Holt, Rinehart & Winston.

Schmenk, B. (2004). Language learning: A feminine domain? The role of stereotyping in constructing gendered learner identities. *TESOL Quarterly, 38*(3), 514–524. https://doi.org/10.2307/3588352

Schunk, D. H., & DiBenedetto, M. K. (2020). Motivation and social cognitive theory. *Contemporary Educational Psychology, 60*, 101832. https://doi.org/10.1016/j.cedpsych.2019.101832

Schunk, D. H., & Usher, E. L. (2019). Social cognitive theory and motivation. In R. M. Ryan (Ed.), *The Oxford handbook of human motivation* (pp. 11-26). Oxford University Press.

Spearman, J., & Watt, H. M. (2013). Perception shapes experience: The influence of actual and perceived classroom environment dimensions on girls' motivations for science. *Learning Environments Research*, *16*(2), 217–238. https://doi.org/10.1007/s10984-013-9129-7

Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, *42*(2), 363–387. https://doi.org/10.1017/S0272263118000311

Tellhed, U., Bäckström, M., & Björklund, F. (2017). Will I fit in and do well? The importance of social belongingness and self-efficacy for explaining gender differences in interest in STEM and HEED majors. *Sex Roles*, *77*(1–2), 86–96. https://doi.org/10.1007/s11199-016-0694-y

Tiedemann, J. (2002). Teachers' gender stereotypes as determinants of teacher perceptions in elementary school mathematics. *Educational Studies in Mathematics*, *50*(1), 49–62. https://doi.org/10.1023/A:1020518104346

Üner, A., Mouratidis, A., & Kalender, İ. (2020). Study efforts, learning strategies and test anxiety when striving for language competence: The role of utility value, self-efficacy, and reasons for learning English. *Educational Psychology*, *40*(6), 781–799. https://doi.org/10.1080/01443410.2020.1736989

Usher, E. L., & Pajares, F. (2008). Sources of self-efficacy in school: Critical review of the literature and future directions. *Review of Educational Research*, *78*(4), 751–796. https://doi.org/10.3102/0034654308321456

Usher, E. L., & Schunk, D. H. (2018). Social cognitive theoretical perspective of self-regulation. In D. H. Schunk & J. A. Greene (Eds.), *Educational psychology handbook series. Handbook of self-regulation of learning and performance* (pp. 19–35). Routledge.

van der Vleuten, M., Jaspers, E., Maas, I., & van der Lippe, T. (2016). Boys' and girls' educational choices in secondary education. The role of gender ideology. *Educational Studies*, *42*(2), 181–200. https://doi.org/10.1080/03055698.2016.1160821

Voyer, D., & Voyer, S. D. (2014). Gender differences in scholastic achievement: A meta-analysis. *Psychological Bulletin*, *140*(4), 1174. https://doi.org/10.1037/a0036620

Wagner, W., Göllner, R., Werth, S., Voss, T., Schmitz, B., & Trautwein, U. (2016). Student and teacher ratings of instructional quality: Consistency of ratings over time, agreement, and predictive power. *Journal of Educational Psychology*, *108*(5), 705-721. https://doi.org/10.1037/edu0000075

Wang, J., & Wang, X. (2019). *Structural equation modeling: Applications using Mplus*. John Wiley & Sons..

West, R., Guven, A., Parry, J. & Ergenekon, T. (2015). *The state of English in higher education in Turkey. A baseline study.* Ankara, Turkey: British Council. Retrieved from http://www.britishcouncil.org.tr/sites/default/files/he_baseline_study_book_web_-_son.pdf.

Wolter, I., Braun, E., & Hannover, B. (2015). Reading is for girls!? The negative impact of preschool teachers' traditional gender role attitudes on boys' reading related motivation and skills. *Frontiers in Psychology*, *6*, 1267. https://doi.org/10.3389/fpsyg.2015.01267

Wu, J.-Y., & Kwok, O. (2012). Using SEM to analyze complex survey data: A comparison between design-based single-level and model-based multilevel approaches. *Structural Equation Modeling: A Multidisciplinary Journal*, *19*(1), 16–35. https://doi.org/10.1080/10705511.2012.634703

Zhang, X. (2019). Foreign language anxiety and foreign language performance: A meta-analysis. *The Modern Language Journal*, *103*, 763–781. https://doi.org/10.1111/modl.12590

***Table 1***
*Model Fit Indices for Confirmatory Factor Analyses*

| | $\chi^2$ | RMSEA | SRMR | CFI | TLI | Factor Loadings |
|---|---|---|---|---|---|---|
| QGSL – Part 1 | 4.68(5) | < .001 | .015 | 1.000 | 1.000 | .51 – .66 |
| QGSL – Part 2 | 9.55(5) | .039 | .022 | .987 | .973 | .50 – .72 |
| MLCAS | 481.28(241)*** | .041 | .030 | .975 | .955 | .52 – .77 |
| QSLL | 27.64(23) | .19 | .016 | .998 | .995 | .45 – .69 |

*** $p < .001$.

**Table 2**
*Descriptive Statistics for Gender Stereotypes, EFL Self-Efficacy, EFL Anxiety and EFL attainment by Gender*

| | Gender | Observed Range | Mean | *SD* | ω | Skewness | Kurtosis | Factor Loadings |
|---|---|---|---|---|---|---|---|---|
| Learners' Stereotypes | Women | 1 – 5 | 2.74 | 0.44 | 0.73 | -0.99 | 2.09 | .49 – .66 |
| | Men | 1 – 5 | 2.77 | 0.51 | 0.77 | 0.06 | 4.57 | .51 – .68 |
| Perceived Teachers' Stereotypes | Women | 1 – 5 | 2.79 | 0.42 | 0.75 | -0.54 | 1.97 | .47–.72 |
| | Men | 1 – 5 | 2.88 | 0.44 | 0.81 | -0.37 | 4.58 | .63 – .79 |
| EFL Self-Efficacy | Women | 1 – 5 | 3.03 | 0.63 | 0.86 | 0.06 | 0.48 | .41 – .63 |
| | Men | 1 – 5 | 3.18 | 0.70 | 0.90 | 0.08 | 0.29 | .46 – .76 |
| EFL Anxiety | Women | 1 – 5 | 2.75 | 0.75 | 0.95 | 0.07 | 0.13 | .47 – .72 |
| | Men | 1 – 5 | 2.43 | 0.81 | 0.96 | 0.15 | -0.21 | .59 – .77 |
| EFL attainment | Women | 27 – 96 | 69.80 | 10.33 | — | -0.55 | 0.96 | — |
| | Men | 30 – 93 | 66.07 | 12.84 | — | -0.78 | 0.44 | — |

**Table 3**

*Pearson's Bivariate Correlations between Gender Stereotypes, EFL Self-Efficacy, EFL Anxiety and EFL Attainment*

| | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1. Learners' Stereotypes | — | .455** | .116 | -.088 | .105 |
| 2. Perceptions of Teacher Stereotypes | .507** | — | .096 | -.134* | .096 |
| 3. EFL Self-Efficacy | -.141* | -.227** | — | -.570** | .316** |
| 4. EFL Anxiety | .071 | .246** | -.535** | — | -.284** |
| 5. EFL Attainment | -.045 | -.065 | .363** | -.324** | — |

*$p < .05$. **$p < .01$. ***$p < .001$.

*Note:* Women below the diagonal, men above.

**Table 4**
*Tests of Measurement Invariance*

| | $\chi^2$ | RMSEA | SRMR | CFI | TLI | Δ RMSEA | ΔCFL | ΔTLI |
|---|---|---|---|---|---|---|---|---|
| **QGSL (Learners' Stereotypes)** | | | | | | | | |
| Configural | 15.255(10) | .041 | .028 | .987 | .973 | | | |
| Metric Invariance | 20.085(14) | .038 | .048 | .985 | .978 | -.003 | -.002 | +.005 |
| Scalar Invariance | 25.361(18) | .037 | .049 | .981 | .979 | -.001 | -.004 | +.001 |
| Residual Invariance | 27.979(23) | .027 | .069 | .987 | .989 | -.010 | +.006 | +.010 |
| **QGSL (Perceptions of Teacher Stereotypes)** | | | | | | | | |
| Configural | 18.267(10)* | .052 | .029 | .978 | .956 | | | |
| Metric Invariance | 25.384(14)* | .052 | .084 | .970 | .957 | <.001 | -.008 | +.001 |
| Scalar Invariance | 33.945(18)* | .054 | .083 | .958 | .953 | +.002 | -.012 | -.004 |
| Partial Scalar Invariance[a] | 29.980 (17)* | .050 | .091 | .966 | .959 | -.004 | +.008 | +.006 |
| **QSLL (EFL Self-efficacy)** | | | | | | | | |
| Configural | 109.271(66)*** | .047 | .038 | .979 | .965 | | | |
| Metric Invariance | 115.442(75)** | .043 | .042 | .981 | .971 | -.004 | +.002 | +.006 |
| Scalar Invariance | 139.502(84)*** | .048 | .050 | .973 | .965 | +.005 | -.008 | -.006 |
| Residual Invariance | 153.648(95)*** | .046 | .057 | .972 | .967 | -.002 | -.001 | +.002 |
| **MLCAS (EFL anxiety)** | | | | | | | | |
| Configural | 741.494(482)*** | .043 | .036 | .973 | .952 | | | |
| Metric Invariance | 804.455(536)*** | .043 | .043 | .972 | .955 | <.001 | -.001 | +.003 |
| Scalar Invariance | 841.471(560)*** | .041 | .045 | .971 | .955 | +.002 | -.001 | <.001 |
| Residual Invariance | 894.688(590)*** | .042 | .050 | .969 | .954 | +.001 | -.002 | <.001 |

*$p < .05$. **$p < .01$. ***$p < .001$.
*Note*. [a] Equality constraint relaxed on item 2 'Please indicate which gender is eager and motivated to learn English according to your English teachers'

**Table 5**

*Results from the Structural Equation Models Predicting EFL Attainment*

| | Learners' gender stereotypes | | | | Learner Perceptions of Teacher Stereotypes | | | |
|---|---|---|---|---|---|---|---|---|
| **Direct Paths** | **β** | ***SE*** | **95% CI** | | **β** | ***SE*** | **95% CI** | |
| | | | **Lower** | **Upper** | | | **Lower** | **Upper** |
| Stereotypes (Women)→ Attainment | .039 | .055 | -.052 | .129 | .067 | .042 | -.002 | .136 |
| Stereotypes (Men) → Attainment | .075 | .042 | .006 | .145 | .033 | .049 | -.047 | .114 |
| Stereotypes (Women)→ Anxiety | .060 | .097 | -.099 | .219 | .253*** | .040 | .187 | .319 |
| Stereotypes (Men) → Anxiety | -.093 | .099 | -.256 | .070 | -.162 | .102 | -.329 | .005 |
| Stereotypes (Women) → Self-efficacy | -.208*** | .044 | -.281 | -.135 | -.279*** | .069 | -.392 | -.166 |
| Stereotypes (Men) → Self-efficacy | .136** | .047 | .058 | .214 | .137 | .107 | -.038 | .313 |
| Anxiety (Women) → Attainment | -.076 | .093 | -.228 | .077 | -.093 | .086 | -.234 | .048 |
| Anxiety (Men) → Attainment | -.061 | .090 | -.208 | .087 | -.058 | .087 | -.200 | .085 |
| Self-efficacy (Women) → Attainment | .412*** | .061 | .313 | .512 | .412*** | .052 | .327 | .497 |
| Self-efficacy (Men) → Attainment | .374* | .198 | .048 | .701 | .383* | .194 | .063 | .702 |
| **Indirect Paths** | | | | | | | | |
| Stereotypes (Women) → Anxiety → Attainment | -.005 | .012 | -.025 | .016 | -.024 | .020 | -.057 | .010 |
| Stereotypes (Men) → Anxiety → Attainment | .006 | .009 | -.008 | .020 | .009 | .011 | -.009 | .028 |
| Stereotypes (Women) → Self-efficacy → Attainment | -.086** | .027 | -.130 | -.042 | -.115*** | .033 | -.168 | -.061 |
| Stereotypes (Men) → Self-efficacy → Attainment | .051** | .015 | .026 | .075 | .053 | .061 | -.048 | .153 |
| **Total Indirect Paths** | | | | | | | | |
| Stereotypes (Women) → Anxiety + Self-efficacy → Attainment | -.090*** | .018 | -.120 | -.061 | -.138*** | .013 | -.160 | -.116 |
| Stereotypes (Men) → Anxiety + Self-efficacy → Attainment | .057*** | .014 | .033 | .080 | .062 | .054 | -.026 | .150 |

*$p < .05$. **$p < .01$. ***$p < .001$.

**Figure 1**

*The Hypothesised Model to Examine the Link Between Learners' Gender Stereotypes and EFL Attainment Through Anxiety and Self-Efficacy*

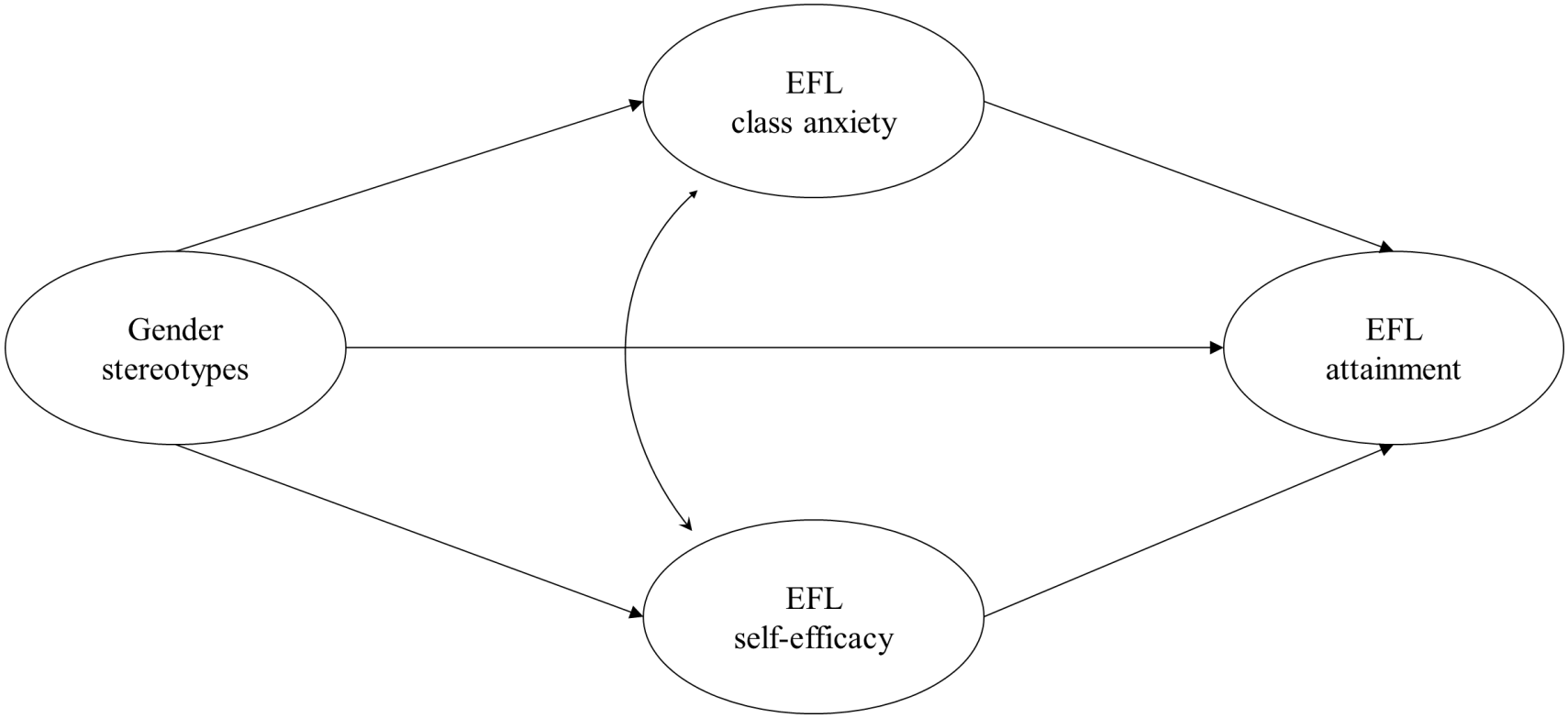
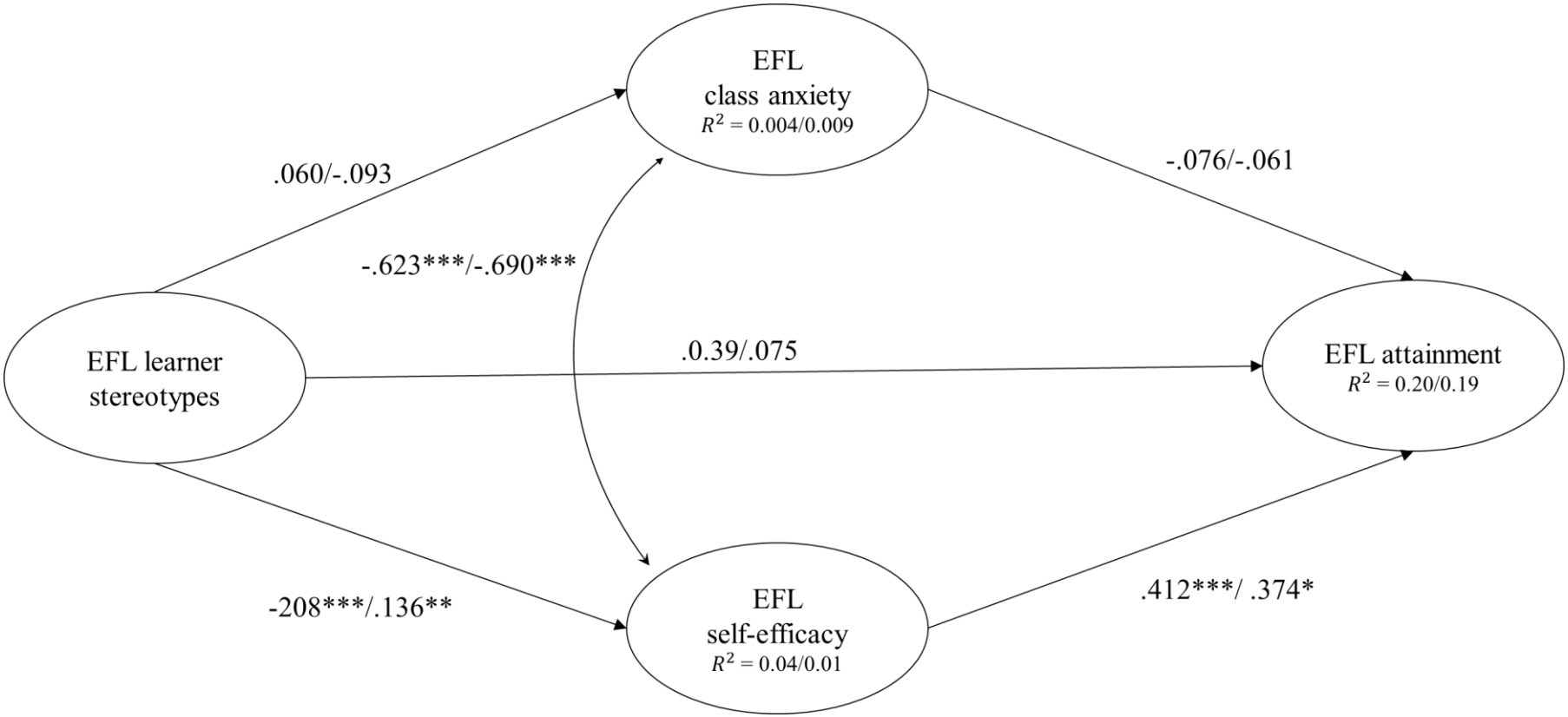


**Figure 2**

*Standardised Path Coefficients to Examine Whether Anxiety and Self-Efficacy Mediated the Relations Between Learners' Gender Stereotypes and EFL Attainment*

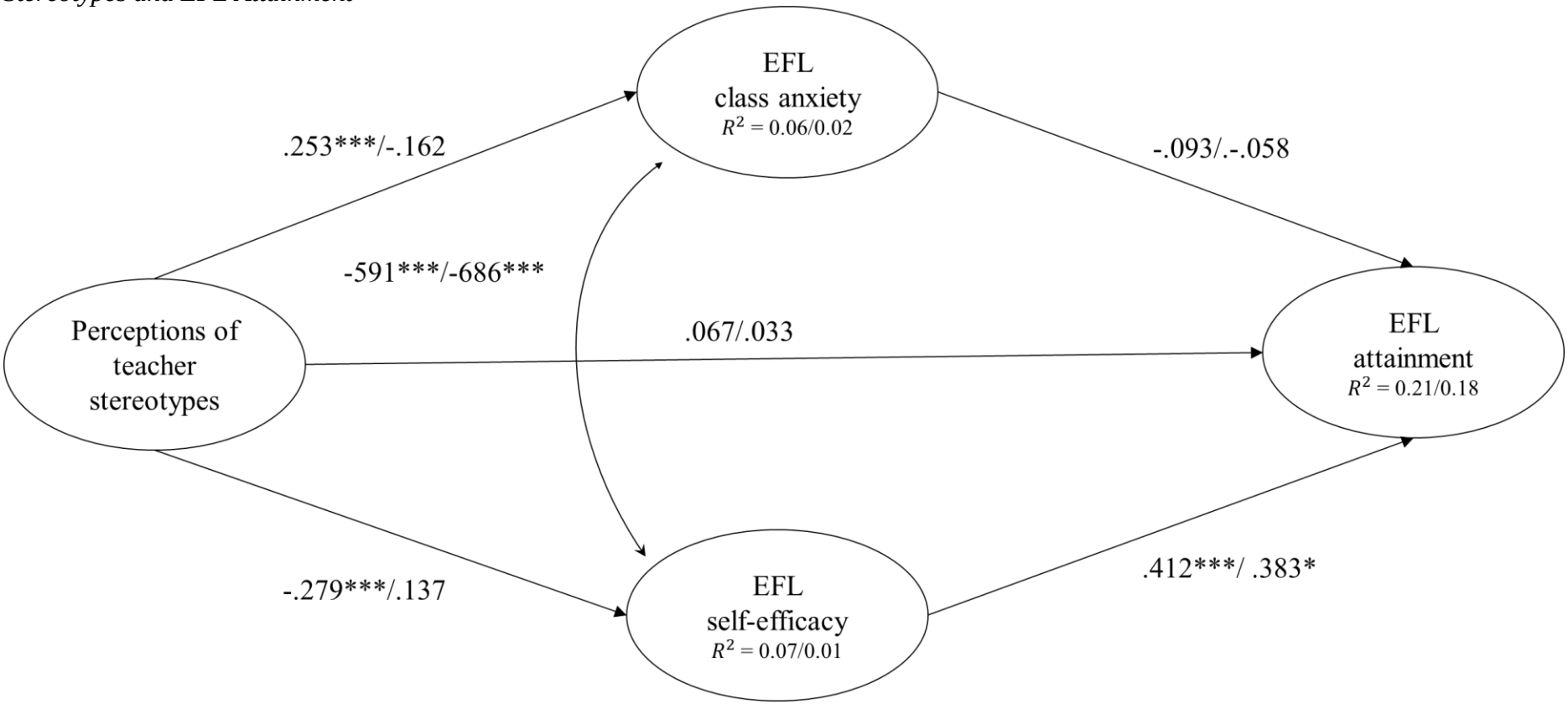



* $p < .05$. ** $p < .01$. *** $p < .001$.

*Note*. Coefficients for females before the dash and males after.

**Figure 3**

*Standardised Path Coefficients to Examine Whether Anxiety And Self-Efficacy Mediated the Relations Between Learner Perceptions of Teacher Stereotypes and EFL Attainment*




* $p < .05$. *** $p < .001$.

*Note*. Coefficients for females before the dash and males after.