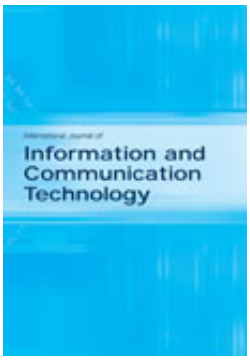


International Journal of Information and Communication Technology

ISSN online: 1741-8070 - ISSN print: 1466-6642
https://www.inderscience.com/ijict

EFL anxiety factor analysis and teaching performance evaluation using R statistical language

Ming Li

DOI: 10.1504/IJICT.2025.10070772

Article History:

Received:	10 March 2025
Last revised:	22 March 2025
Accepted:	22 March 2025
Published online:	08 May 2025

EFL anxiety factor analysis and teaching performance evaluation using R statistical language

Ming Li

School of Humanities,
Ningbo University of Finance and Economics,
Ningbo 315175, China
Email: lm25406@163.com

Abstract: With the popularity of English education, students' anxiety has an important impact on learning effect. The purpose of this study is to explore the anxiety factors of English learners and their impact on teaching performance, and use R language for data analysis. Through the questionnaire survey of English majors' anxiety, this paper collected the quantitative data of language anxiety, test anxiety, communication anxiety and other anxiety factors. On this basis, the factor analysis method is used to extract and verify the English learning anxiety factors, and determine the main anxiety factors and their internal structure. At the same time, the relationship between anxiety factors and academic performance is discussed by using correlation analysis and regression analysis. The results show that language anxiety and communication anxiety are the main sources of students' anxiety, and they are significantly negatively correlated with students' teaching performance.

Keywords: EFL; anxiety factor; R language; factor analysis; teaching achievement.

Reference to this paper should be made as follows: Li, M. (2025) 'EFL anxiety factor analysis and teaching performance evaluation using R statistical language', *Int. J. Information and Communication Technology*, Vol. 26, No. 12, pp.70–88.

Biographical notes: Ming Li received her PhD degree from Department of English of National Chengchi University in Taiwan. Currently, she works as a Lecturer in Ningbo University of Finance and Economics. Her research interests include English education, and computer-assisted instruction (CAI).

1 Introduction

With the acceleration of globalisation, English, as one of the most widely used languages in the world, has become an indispensable subject in the education system of various countries. For students from non-native English-speaking countries, learning English as a foreign language (EFL) faces many challenges. Studies have found that foreign language learners often have anxiety in the process of learning, which not only affects learning motivation, but also may directly interfere with the effect of language learning, and even affect academic performance (Horwitz, 2001). Therefore, anxiety in foreign language learning has gradually become a hot topic in educational research, especially in EFL learning environment.

Anxiety is an emotional reaction that individuals produce when facing threats or uncertain situations. It is usually accompanied by physiological reactions (such as tachycardia and shortness of breath) and cognitive negative evaluations (such as worry and fear). In the process of foreign language learning, the sources of anxiety are complex and diverse, mainly including language anxiety, communication anxiety, test anxiety and other dimensions (Tran, 2011). These anxiety factors are usually intertwined, and will have different degrees of influence on students' language learning performance.

Language anxiety refers to the anxiety caused by students' doubts and self-confidence about their language ability in the process of language learning. In the process of learning language skills such as listening, speaking, reading and writing, many EFL learners tend to feel inferior and worry that they cannot accurately express their thoughts or understand others (Liu and Huang, 2011). This anxiety not only limits the improvement of students' language skills, but also may cause them to behave negatively in class, avoid participating in language activities, and further aggravate learning difficulties.

Communication anxiety is mainly manifested in students' tension and fear when communicating with others in a foreign language (Zwingmann et al., 2017). This anxiety usually appears in public speaking, class discussion or communication with strangers. EFL learners tend to worry too much about the accuracy and fluency of language expression in communication situations, which leads to communication barriers. Communication anxiety is particularly significant in the early stage of language acquisition, especially in the environment without adequate language practice, students' anxiety often has a negative impact on their language output.

Test anxiety refers to the anxiety of students in the face of language tests. EFL learners usually pay close attention to test scores, especially in standardised tests, and this anxiety is more prominent. Test anxiety may lead to students' inability to give full play to their language ability in examinations, especially in listening and speaking. Test anxiety will make students perform poorly under time constraints and scoring pressure, and then affect their academic performance (Spielberger et al., 2015).

The effect of anxiety on EFL achievement has been verified by a large number of empirical studies. Many studies have found that there is a significant negative correlation between EFL learners' anxiety and academic performance (Hanifa, 2018). Specifically, anxiety may affect students' learning effect through a variety of ways. First of all, anxiety may reduce students' self-confidence and enthusiasm, leading to their lack of motivation in the learning process, and then affect their learning effectiveness. Secondly, anxiety may also reduce students' learning efficiency in class by causing distraction and memory loss. For example, test anxiety may make it difficult for students to concentrate during the test, thus affecting their performance. Finally, anxiety may also lead to students' negative emotional reactions, such as fear, irritability, etc. these negative emotions will aggravate anxiety and make students fall into a vicious circle.

However, the impact of anxiety is not one-way. Some studies also pointed out that moderate anxiety may instead become a driving force to help students maintain their sense of urgency and concentration in learning (Yan and Horwitz, 2008). Therefore, the effect of anxiety is not completely negative, and its degree of influence is closely related to students' individual differences, learning environment and teachers' educational strategies.

Although the researches on anxiety in EFL learning have been increasing in recent years, most of them focus on qualitative analysis and questionnaire survey, lacking systematic quantitative analysis of anxiety factors. At the same time, most of the existing

studies focus on a single anxiety factor, such as language anxiety or test anxiety, while few studies consider the comprehensive impact of multiple anxiety factors on students' academic performance. In addition, although previous studies have shown that there is a negative correlation between anxiety and academic achievement, the specific mechanism between anxiety factors and student achievement has not been systematically discussed, especially how to alleviate students' anxiety and improve learning effect through scientific teaching methods, which still needs further research.

1.1 Related work

Anxiety refers to an individual's emotional response to future uncertainty and potential threats, usually accompanied by physiological reactions (such as rapid heartbeat and shortness of breath) and psychological reactions (such as worry and fear). In EFL learning, the definition of anxiety is diverse. Luo (2013) believes that foreign language anxiety refers to students' tension, anxiety and fear in the process of foreign language learning, which may hinder the effectiveness of their language learning. According to the forms and sources of anxiety, anxiety in EFL learning can be divided into three types: language anxiety, communication anxiety and test anxiety.

Language anxiety refers to the anxiety caused by students' doubts about language ability or low self-evaluation when using a foreign language (Svalberg, 2012). This anxiety is usually manifested in students' fear of inaccurate language expression, grammatical errors or non-standard pronunciation. Communication anxiety refers to students' anxiety in the process of actual language communication, especially in oral communication with others. This anxiety stems from the fear of others' evaluation, which is usually manifested in the fear of expressing their ideas in public (Jiang and Dewaele, 2019). Test anxiety is a kind of anxiety caused by students' worry and fear about the test results when facing the language test (Bandelow et al., 2017). This anxiety may affect students' performance, especially in language tests with high time pressure.

A large number of studies have shown that EFL learners' anxiety has a significant negative impact on their academic performance. Pekrun et al. (2023) proposed that anxiety not only affects students' language input and output ability, but also may lead to their negative emotions and low self-confidence in the classroom, thus affecting their academic performance. Russell (2020) found that language anxiety is closely related to students' oral ability and listening comprehension. Students with strong anxiety usually perform poorly in oral communication. Hashemi's (2011) input hypothesis theory points out that anxiety will hinder the effective reception of language input, thus affecting language learning. The research shows that communicative anxiety mainly affects students' oral and interactive abilities. Fuller et al. (2016) believe that students with communication anxiety tend to avoid oral communication and collective activities, which will affect their language output and the development of oral ability. Vonder-Embse et al. (2018) found that test anxiety not only affects students' test performance, but also may have a negative impact on students' self-efficacy. Szafranski et al. (2012) research shows that test anxiety is usually related to students' learning pressure, emotional fluctuations and distraction, thus reducing their academic performance.

EFL learners' anxiety is affected by many factors, including individual differences, learning environment, cultural background and teachers' teaching methods.

According to the research, differences in individual personality characteristics, extroversion and introversion have a significant impact on the level of anxiety (Bados

et al., 2010). Extroverts usually show a lower level of anxiety, while introverts are more likely to have anxiety. EFL learning environment has an important impact on anxiety. Nielsen and Harder (2013) believed that the classroom atmosphere, the relationship between teachers and students and the interaction between students all had an impact on students' anxiety. In a supportive and interactive learning environment, students' anxiety is often light. Cultural background also has a significant impact on EFL learners' anxiety. Due to the high sensitivity to failure in the cultural tradition, Chinese students usually show a strong sense of anxiety in foreign language learning (Diep et al., 2022). In contrast, Western students show more positive challenge attitude. Teachers' teaching strategies, classroom management methods and evaluation criteria will also affect students' anxiety level. Research shows that the teaching method with strong interaction and good emotional support can effectively reduce students' anxiety (Vanden-Berghe et al., 2018).

Anxiety not only affects students' academic achievement, but also indirectly affects their academic achievement by affecting their learning motivation. Liu (2021) proposed that anxiety may reduce students' learning motivation, especially in foreign language learning.

It is found that anxiety is closely related to learners' intrinsic motivation. High anxiety students usually lack intrinsic motivation and are often lack of interest in foreign language learning, and pay more attention to test scores and external evaluation. Contrary to intrinsic motivation, students with extrinsic motivation usually show stronger coping strategies when facing anxiety, such as overcoming anxiety by studying hard and seeking teachers' help (Fiddiyasari and Pustika, 2021).

Effective anxiety management strategies are the key to alleviate EFL learners' anxiety. Research shows that learners' coping strategies, emotion regulation ability and social support system all play an important role in anxiety management.

Emotion regulation theory points out that students can reduce the negative impact of anxiety on learning through effective emotion regulation (Vela et al., 2022). The support of teachers and classmates is very important to alleviate anxiety. A supportive learning environment can help students build self-confidence and reduce the negative effects of anxiety (Diamond and Alley, 2022).

Through the cognitive reconstruction technology in cognitive behavioural therapy, students can change their negative views on anxiety sources, so as to effectively reduce anxiety (Li et al., 2022).

Although the research on anxiety in EFL learning has made important progress, there are still some limitations. First of all, the existing studies mostly use the questionnaire survey method, and lack of dynamic observation of anxiety. Secondly, there are few studies on the interaction between different anxiety factors and their comprehensive impact on academic performance. Future research can pay more attention to the multidimensional analysis of anxiety, and use the longitudinal research design to explore the causal relationship between anxiety and academic achievement.

1.2 Contribute

In view of the shortcomings of the above research, this study aims to collect EFL learners' anxiety data through a questionnaire survey and use R language for data analysis to systematically explore EFL learners' anxiety factors and their relationship with teaching achievement. Specifically, this study has the following two main purposes.

Quantitative analysis of EFL anxiety factors: through factor analysis, the main anxiety factors of EFL learners are extracted, and the internal relationship between different anxiety factors is investigated. By using R language for data processing, this paper further reveals the mechanism of anxiety factors in students' learning process.

To explore the relationship between anxiety factors and teaching performance: combined with students' final examination results, this paper analyses the influence of anxiety factors on teaching performance, and discusses how different anxiety factors interact and jointly affect students' academic performance.

This study not only helps to deeply understand the multidimensional characteristics of EFL learners' anxiety and its impact on academic performance, but also provides a new idea for teachers to help them identify students' anxiety sources, adopt more targeted teaching strategies, improve the learning environment, so as to alleviate students' anxiety and improve teaching effect.

2 Relevant theoretical foundations

2.1 Theoretical framework of anxiety

As an emotional response, anxiety usually involves cognitive, emotional and physiological processes. In EFL learning, anxiety is not only the result of students' individual emotional experience, but also closely related to cognitive and social interaction in the learning process. In order to fully understand the anxiety in EFL learning, we need to explore the formation mechanism of anxiety in different theoretical frameworks, and explain its internal relationship combined with mathematical models.

The cognitive affective theoretical framework focuses on the cognitive and affective processes of anxiety. According to the cognitive affective model of MacIntyre and Gardner (1991), anxiety is an emotional response caused by an individual's cognitive assessment of a task.

In the cognitive affective model, the occurrence of anxiety depends on students' cognitive assessment of learning tasks, which is composed of the following factors.

1 Self-efficacy: Learners' belief in whether they can successfully complete language tasks.

2 External evaluation: Learners' concerns about external evaluation such as teachers, peers and examinations.

3 The gap between expectation and reality: the difference between students' expectation and actual performance.

Anxiety is usually exacerbated by low self-efficacy, excessive concern about external evaluation, and the discrepancy between expectations and actual performance. We can use a simple cognitive assessment formula to express the generation of anxiety:

$$A = f(E, R, S) \tag{1}$$

where A is the anxiety level, E is the expectation (students' expectation of task completion), R is the actual performance (students' ability to actually complete the task), and S is the students' self-efficacy (that is, students' evaluation of their ability). When the

current expectation is greater than the actual performance, and the self-efficacy is low, the sense of anxiety will increase.

The proposed state trait anxiety theory is another important framework to explain anxiety. According to this theory, anxiety can be divided into two types. The first is state anxiety, which refers to the temporary anxiety of individuals in a specific situation. It is influenced by external situations (such as examinations, oral tests, etc.) and is usually short-lived. The second is trait anxiety, which refers to the anxiety tendency that individuals generally have in daily life, and reflects the anxiety sensitivity of individuals to different situations.

State anxiety and trait anxiety are interrelated. Mathematically, we can describe the relationship between state anxiety and trait anxiety through the following formula:

$$A_{state} = \alpha(A_{trait}) + \beta \tag{2}$$

where A_{state} is the level of state anxiety and A_{trait} is the level of individual trait anxiety; α and β are parameters related to individual characteristics (e.g., individual emotion regulation ability, adaptability to anxiety, etc.).

The theoretical framework of social psychology of anxiety mainly focuses on the social causes of anxiety, especially learners' social interaction, cultural background and sensitivity to others' evaluation. EFL learners' anxiety is not only affected by their individual characteristics, but also by social factors such as learning environment, teacher expectations, peer evaluation and so on.

Learners are particularly sensitive to the evaluation of others (such as teachers, classmates, examination judges, etc.). This sensitivity will lead to students' anxiety in the process of language learning, especially when they need to express or interact openly. Students from different cultural backgrounds have different perceptions of anxiety and coping styles. For example, learners of collectivist culture may feel greater pressure of social evaluation when expressing in public, resulting in strong anxiety. The anxiety model under the framework of social psychology can be expressed by the following formula:

$$A = f(S, C) \tag{3}$$

where A represents the anxiety level, S represents the sensitivity of students to social evaluation (such as the degree of concern for teacher evaluation and classmate evaluation), and C represents the influence of cultural background and social interaction (such as the constraint of culture on individual expression and sensitivity to failure).

Anxiety is not only a cognitive and psychosocial response, but also accompanied by physiological changes. For example, anxiety may lead to physiological reactions such as rapid heartbeat, shortness of breath, sweating, etc. This process is regulated by the autonomic nervous system and is part of the physiological emotional response model.

According to Yerkes Dodson rule, moderate anxiety can improve learning performance, but excessive anxiety can lead to performance decline. The relationship between physiological response and emotional response can be expressed by the following formula:

$$P = \alpha(A) - \beta(A^2) \tag{4}$$

where P stands for academic performance (such as test scores), A stands for anxiety level, α and β are constants related to anxiety intensity and learning performance. According to

Yerkes Dodson rule, anxiety will promote learning performance when it is moderate, but when the anxiety level is too high (that is, when A is large), it will have a negative impact on learning performance (that is, P decreases).

The emotion regulation model of anxiety focuses on how individuals deal with anxiety through emotion regulation strategies. The level of emotion regulation ability determines the influence of anxiety on academic performance. Individuals can regulate anxiety through different regulation strategies (such as cognitive reconstruction, emotional inhibition, etc.). The effect of emotion regulation can be expressed by the following formula:

$$A' = A - \gamma(R) \tag{5}$$

where A' is the adjusted anxiety level, A is the original anxiety level, and R is the effect of emotion regulation strategies (such as cognitive reconstruction, emotion expression, etc.), γ is the coefficient of regulation effect, indicating the effectiveness of emotion regulation strategies. Through effective emotion regulation strategies, students can reduce anxiety and improve learning performance.

2.2 Factor analysis theory

Factor analysis is a commonly used statistical method, which is widely used in psychology, pedagogy, sociology and other fields to study the potential relationship between a group of observed variables. Through factor analysis, researchers can identify the potential structure hidden behind a large number of observation data, understand the internal relationship between variables, and explain the changes of data through factors.

The goal of factor analysis is to express a set of relevant observed variables as a linear combination of a few potential factors, taking into account the measurement error. Through factor analysis, researchers can find the internal relationship between variables and reduce the data dimension, so as to facilitate understanding and interpretation. The basic model of factor analysis usually adopts the following forms:

$$X_i = \lambda_{i1}F_1 + \lambda_{i2}F_2 + \ldots + \lambda_{im}F_m + \varepsilon_i \tag{6}$$

where X_i is the i^{th} observation variable (e.g., different anxiety dimensions in EFL learning, language anxiety, communication anxiety, etc.), F_j is the j^{th} potential factor (e.g., potential anxiety factor), λ_{im} is the factor load, indicating the degree to which the factor explains the observation variable, and ε_i is the observation error term, representing the changes caused by factors other than potential factors.

Factor analysis usually includes the following main steps:

1 Data preparation.

   Before factor analysis, it is necessary to ensure the applicability and reliability of the data. Common inspection methods include Kaiser Meyer Olkin measure of sampling efficiency, which is used to evaluate whether the sample size is sufficient. The closer the KMO value is to 1, the data is suitable for factor analysis. Bartlett's test of sphericity is also used to test whether the data matrix has sufficient correlation and whether factor analysis can be carried out.

2 Extraction factor.

The goal of factor extraction is to select the number of factors that can explain the variation of the original data. The commonly used extraction method is principal component analysis (PCA), which can maximise the variance of the data and extract the most representative factors by projecting the data onto the new coordinate axis (PCA). The maximum likelihood method (MLE) estimates the parameters in the factor model by maximising the likelihood function of the observed data.

3 Rotation factor.

After factor extraction, factor rotation is usually required to improve the interpretability of factor model. The methods of rotation include orthogonal rotation, such as varimax rotation, assuming that the factors are independent. There is also oblique rotation: for example, promax rotation allows correlation between factors. The rotated factor load matrix is usually easier to interpret, and can make the load of each factor on the observation variable more concentrated or more characteristic.

4 Factor score calculation.

After the factor analysis is completed, it is necessary to calculate the scores of each sample on different factors. These scores can be used for subsequent analysis, such as regression analysis or cluster analysis. The factor score is usually calculated by the following formula:

$$F_j = \sum_{i=1}^{n} \lambda_{ij} X_i \tag{7}$$

where F_j is the score of the j^{th} factor, λ_{ij} is the factor load, and X_i is the value of the i^{th} observed variable.

2.3 Statistical models and methods

In this study, we use factor analysis to identify EFL learners' anxiety factors, and use R language for data processing. R language provides a variety of packages and functions to implement factor analysis. Commonly used packages include psych, factoextra, lavaan, etc.

Factor analysis methods commonly used in R language include PCA and maximum likelihood estimation (MLE). In R language, we can use the factorial() function for factor analysis: fa_result <- factorial (factors = 3, covmat = covariance_matrix) summary (fa_result), where factors = 3 means to extract three factors. Covmat is the covariance matrix of data, which can be calculated by cov () function. Summary (fa_result) is used to output the results of factor analysis, including information such as factor load and variance interpretation proportion.

In order to explore the relationship between anxiety factors and students' teaching performance, we can use correlation analysis, regression analysis and other methods. In this study, we will use multiple regression analysis to evaluate the impact of anxiety factors on teaching performance. The expression of multiple regression model is:

$$Y = \beta_0 + \beta_1 F_1 + \beta_2 F_2 + \ldots + \beta_m F_m + \varepsilon \tag{8}$$

where Y is the teaching score (the final exam score on schedule), and F_i is the extracted anxiety factor, β_0 is the regression intercept, β_i is the regression coefficient, indicating the impact of each anxiety factor on teaching performance, ε is the error term.

In R language, we can use lm() function for regression analysis, model <- lm(score ~ factor1 + factor2 + factor3, data = data) summary(model).

In order to further verify the relationship between anxiety factors and students' teaching performance, structural equation model (SEM) can be used. SEM can simultaneously consider multiple factors and their relationships, and provide more accurate model fitting and path analysis. The commonly used structural equation model package in R language is lavaan. Through SEM model, we can more intuitively understand the path relationship between anxiety factors and teaching achievement.

3 Analysis of EFL anxiety factors based on R language

3.1 Research design

This study uses a cross-sectional design to analyse the relationship between EFL anxiety and academic performance by collecting student data at a specific time point. The main steps of the study include:

1 Measurement of anxiety factors. Through the quantitative questionnaire survey, the anxiety feeling of students in EFL learning was collected, and the potential anxiety factors were extracted by factor analysis.

2 Evaluation of teaching achievements. Students' midterm or final grades are used as a measure of their English learning performance.

3 Data analysis. Factor analysis was used to extract anxiety factors, and regression analysis was used to explore the relationship between anxiety factors and academic performance.

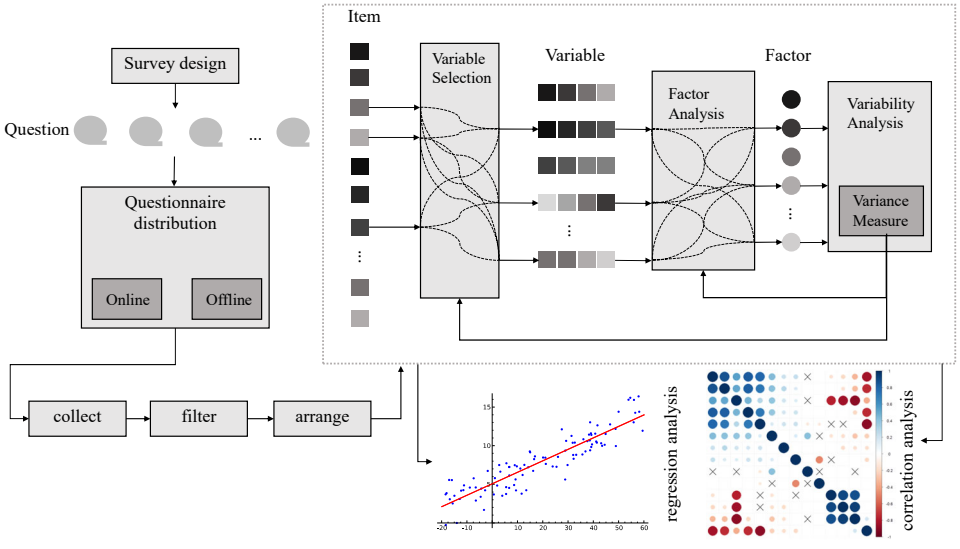
The research framework is shown in Figure 1.

3.2 Data collection and analysis

The sample of this study is from the undergraduate students of EFL course in a university. The subjects of the study are all English majors and non-English majors, covering students of different grades and majors. Convenience sampling was used for sample selection, that is, the individuals who were willing to participate in the survey among the students in school were selected.

The sample size is expected to be 200 to 300 students to ensure the representativeness of data and the effectiveness of statistical analysis. According to the size of the sample and the purpose of the study, it is planned to adopt the multi-stage sampling method and randomly select participants in multiple grades.

Figure 1 Research framework (see online version for colours)



In order to measure the level of anxiety in EFL learning, this study designed a multi-dimensional anxiety questionnaire. The questionnaire is mainly divided into the following parts:

- Personal information: including basic information such as gender, age, grade and years of English study.
- Anxiety scale: using the foreign language classroom anxiety scale (FLCAS) as the core tool, the scale includes language anxiety, communication anxiety, test anxiety and other dimensions. Each dimension includes multiple declarative sentences, and participants choose the corresponding score according to their own feelings (for example, from ‘strongly disagree’ to ‘strongly agree’).
- Self-assessment: it includes students’ self-assessment of their English ability, aiming to understand the relationship between students’ self-confidence and anxiety.
- Academic performance data: Students’ midterm or final exam scores are used as the standard to measure their English academic performance.

FLCAS was proposed by Horwitz (1986) and is widely used to study anxiety in foreign language learning, with good reliability and validity. In this study, the scale has been modified and adjusted to meet the characteristics of EFL learning in the Chinese environment.

Questionnaires are distributed online and offline. First, the electronic version of the questionnaire is released to students through the school’s online learning platform; Then, the research team also distributed paper questionnaires to students of different grades to ensure that they covered different groups of students. Each questionnaire is accompanied by a brief description, emphasising the anonymity of filling in and the confidentiality of the purpose of the survey, so as to ensure that students can truly reflect their own situation.

After receiving the questionnaire, first clean up the data and eliminate the invalid questionnaire (such as too many missing values or inconsistent questionnaires). Code all questionnaire data, use appropriate digital representation to convert questionnaire answers, and ensure the numerical processing of each question and answer. Because the scale ranges of different items are different, all data are standardised (such as Z-score Standardisation) to ensure that the scale dimensions are comparable. Factor analysis was performed using R language or SPSS to extract potential anxiety factors. Specifically, factor extraction is performed through exploratory factor analysis (EFA), and the factor load matrix is optimised using rotation methods such as varimax rotation.

In order to explore the relationship between anxiety factors and students' academic performance, this study uses Pearson correlation analysis and multiple linear regression analysis. Specifically, the anxiety factor as an independent variable, students' academic performance as a dependent variable, test the influence of anxiety factor on academic performance.

Pearson correlation coefficient is used to measure the linear relationship between features and target variables. Suppose there are characteristic x and target variable y, and their correlation is calculated as:

$$Corr(x, y) = \frac{\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \overline{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \overline{y})^2}} \tag{9}$$

where x_i, y_i is the i^{th} sample value. $\overline{x}$, $\overline{y}$ is the mean value of variables X and Y respectively. n is the total number of samples.

Feature redundancy can be evaluated by calculating the correlation between features. If the correlation between the two features is high, there is redundancy between them. The common redundancy evaluation method is to calculate mutual information between features, and its calculation formula is:

$$I(X, Y) = \sum_{x \in X}\sum_{y \in Y} p(x, y) \log \frac{p(x, y)}{p(x)p(y)} \tag{10}$$

where $p(x, y)$ is the joint probability distribution and $p(x)$, $p(y)$ is the marginal probability distribution.

In the form of electronic questionnaire, 303 questionnaires were collected, three abnormal data were eliminated, and 300 valid questionnaires were retained, with an effective rate of about 99%. Among them, 95 were boys and 205 were girls. The obtained data were used to test the consistency of the questionnaire using the alpha function in the psych package of R language, and the Cronbach alpha coefficient was 0.94, which met the internal consistency requirements of questionnaire reliability. On this basis, the statistical analysis tools in R language are used to analyse the factors of English classroom anxiety, and to explore the relationship between anxiety and achievement.

Factor analysis is the core analysis method of this study, which aims to extract potential factors from the anxiety dimension of students' self-evaluation. After data standardisation, EFA was used to extract representative anxiety factors. The factors were extracted by PCA or maximum likelihood analysis (MLE). The appropriate number of factors is usually determined by Kaiser criterion (eigenvalue greater than 1) and screen plot. The extracted factors are rotated using varimax or promax to facilitate factor

interpretation. According to the factor load matrix, name each factor and explain the potential anxiety dimensions.

In the process of data collection and processing, this study strictly followed the ethical norms. All participants volunteered to fill in the questionnaire. All data were kept anonymous and only used for academic research. The personal information obtained during the study is strictly confidential to ensure that it will not cause any infringement on the privacy of participants. In addition, the research team also ensured the fairness and impartiality of the research and avoided any potential bias affecting the research results.

3.3 Evaluation

KMO value is used to measure whether the sample data is suitable for factor analysis. The larger the value, the higher the fitness of factor analysis. KMO values range from 0 to 1, and the formula is as follows:

$$KMO = \frac{\sum_i \sum_j |r_{ij}|}{\sum_i \sum_j |r_{ij}| + \sum_i \sum_j |m_{ij}|} \tag{11}$$

where r_{ij} is the correlation coefficient between variables and m_{ij} is the inverse of the sample covariance matrix.

Bartlett's test of sphericity is used to evaluate whether the correlation of sample data is strong enough to support factor analysis. If the p value is less than 0.05, it indicates that the data is suitable for factor analysis.

$$x^2 = \left(N - 1 - \frac{2p+5}{6} \right) \cdot \left[\sum_i \sum_j \left(r_{ij} \right)^2 \right] \tag{12}$$

where N is the sample size, p is the number of variables, and r_{ij} is the correlation coefficient.

The coefficient of determination R^2 measures the degree to which the regression model explains data variability. Its value is between 0 and 1, and the closer the value is to 1, the higher the model's fit to the data and the stronger its explanatory power.

$$R^2 = 1 - \frac{\sum_{i=1}^{n} \left(Y_i - \hat{Y}_i \right)^2}{\sum_{i=1}^{n} \left(Y_i - \overline{Y}_i \right)^2} \tag{13}$$

where Y_i is the actual observation value, $\hat{Y}_i$ is the regression prediction value, and $\overline{Y}_i$ is the average of all observation values.

The adjusted coefficient of determination takes into account the number of variables in the model and is suitable for multiple regression analysis, which can better measure the fit of the model.

$$\overline{R}^2 = 1 - \left(\frac{(1 - R^2)(n-1)}{n - p - 1} \right) \tag{14}$$

where n is the sample size, and p is the number of independent variables

Standardised regression coefficients (beta coefficients) are used to measure the strength and direction of the influence of the independent variable (anxiety factor) on the dependent variable (student performance). Standardised regression coefficients enable the comparison of the influence of different variables.

$$\hat{\beta} = \frac{Cov(X, Y)}{Var(X)} \tag{15}$$

where Cov (X, Y) is the covariance between the independent variable and the dependent variable, and $Var(X)$ is the variance of the independent variable.

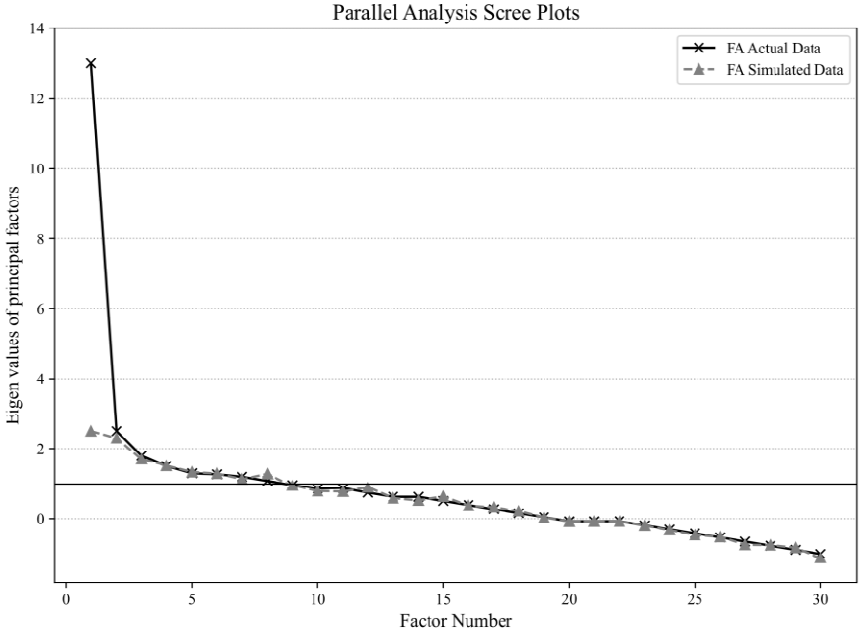
The p-value is used to test whether the regression coefficient is significant. If the p-value is less than 0.05, it indicates that the factor has a significant impact on academic performance.

4 Experiment

4.1 Data adaptability verification

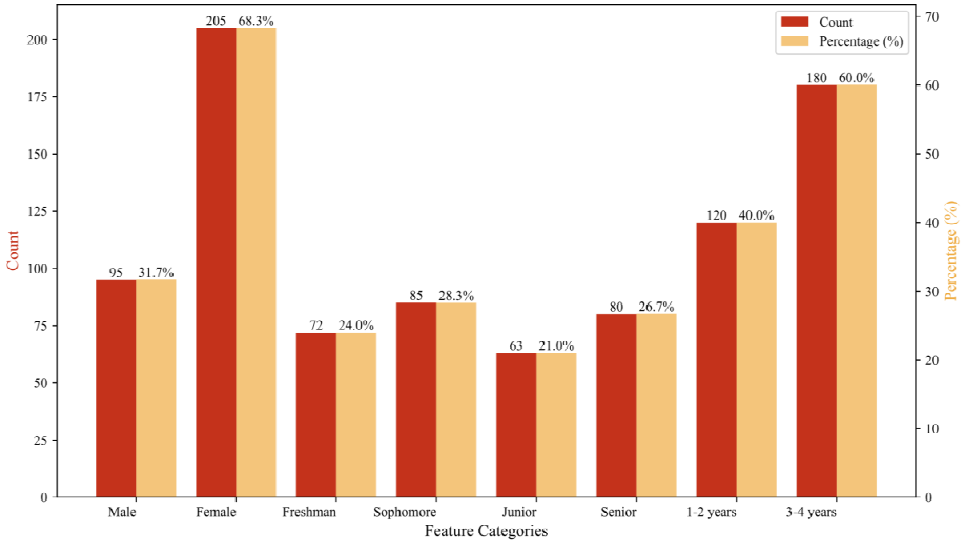
The polychoric function in the psych package of R language is used to transform the data into multi grid relational data. On this basis, KMO and cortest. Bartlett function are used to verify the adaptability of the converted data. The results show that the KMO value is 0.85 > 0.8, and the Bartlett test result is significant (3,920, DF = 529, P = 0.000 < 0.01), which indicates that the data meet the conditions of factor analysis. In the R language's psych package, the cortest. Bartlett() function automatically completes the above calculation process.

Figure 2 Parallel analysis scree plots



The extraction of factor number is of great significance in factor analysis. Use the fa.parallel function in the R language psyche package to iterate the data 100 times and select the factor extraction method for extraction. The results are shown in Figure 2. It can be seen that there are about 6 actual data points distributed above the dotted line of factor fitting data, so it is recommended to extract 6 factors. At the same time, try other methods, the results show that the six factors are more reasonable to explain the variables, so choose six factors to fit.

Figure 3 Distribution of features by count and percentage (see online version for colours)



4.2 Descriptives

The basic data of the participants were analysed by descriptive statistics in order to understand the basic characteristics of the sample. Figure 3 shows the basic situation of the sample, including the frequency distribution of variables such as gender, grade and years of English learning.

Through the descriptive analysis of the problems in the anxiety scale, it shows that most participants show a certain degree of anxiety in EFL learning, especially in oral expression and exam related problems. The average score of anxiety is 3.72, and the standard deviation is 0.85, indicating that students generally have certain anxiety.

4.3 Factor analysis results

In order to extract the anxiety factors in EFL learning, EFA was carried out. According to Kaiser criterion (eigenvalue greater than 1) and screen plot, three factors were determined to be extracted, namely 'language anxiety' (factor 1), 'communication anxiety' (factor 2) and 'test anxiety' (factor 3). The factor load matrix shows that the loads of each factor are as follows:

As can be seen from Figure 4, factor 1 (language anxiety) mainly includes anxiety items of oral English, writing and listening, factor 2 (communication anxiety) mainly

includes anxiety items of classroom participation and communication, and factor 3 (examination anxiety) mainly includes anxiety items related to examination and review. This shows that anxiety in EFL learning can be divided into three main types: language anxiety, communication anxiety and test anxiety.

Figure 4 Factor loadings heatmap (see online version for colours)

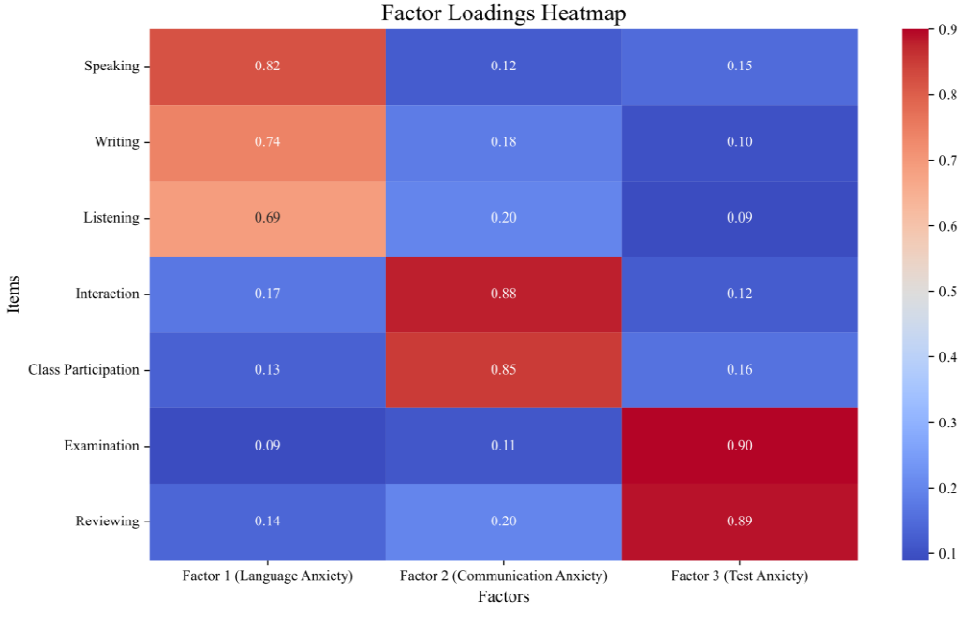
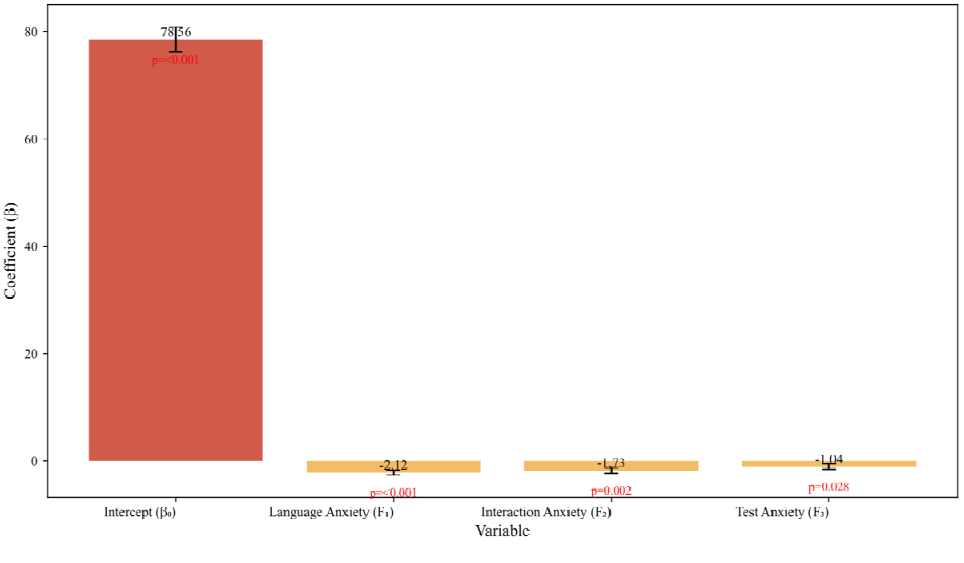


Figure 5 Regression coefficients with standard errors (see online version for colours)



4.4 Regression analysis results

After factor analysis extracted three anxiety factors, multiple regression analysis was carried out to explore the relationship between these anxiety factors and students' academic performance. Students' final exam scores were used as dependent variables, and three anxiety factors were used as independent variables for regression analysis.

The results of regression analysis show that language anxiety has the most significant negative impact on students' performance, followed by communication anxiety, and test anxiety also has a significant negative impact, but relatively small. The coefficient of determination of the overall regression model was 0.35, indicating that anxiety factors could explain 35% of the differences in academic performance.

4.5 Experimental results

This study identified three major anxiety factors in EFL learning through factor analysis: language anxiety, communication anxiety and test anxiety. These factors have a strong influence on EFL learning, especially language anxiety, which shows that students will feel a strong sense of anxiety when learning English, especially in the learning of oral and listening skills. Although communication anxiety and test anxiety also significantly affect students' academic performance, their influence is relatively weak.

This study found that language anxiety has a significant negative impact on students' academic performance, especially in oral and writing tasks. Language anxiety may cause students to feel nervous and inferior when expressing themselves, thus affecting their actual performance.

Communication anxiety is mainly reflected in students' tension in classroom interaction and communication with others. Although this type of anxiety has a certain impact on performance, its negative effect is mild, which may be because students' anxiety about communication is relatively short-term and adjustable, and may not directly affect long-term learning outcomes.

The negative effect of test anxiety is obvious. Although its regression coefficient is small, its effect on academic performance still exists. Students' test anxiety may lead to excessive anxiety during review, which will affect their learning efficiency and test performance.

The results of regression analysis show that the anxiety factors in EFL learning significantly affect students' academic performance. Language anxiety has the most significant negative impact on academic performance, which shows that students' emotional state and psychological pressure directly affect their learning effect in the process of English learning. Because anxiety may affect students' ability of attention, memory and emotion regulation, this explains why students with higher anxiety tend to have lower grades.

Although other factors (such as learning methods, learning motivation, etc.) can also affect students' performance, anxiety, as an emotional state, has proved that its negative effect on learning cannot be ignored. Therefore, this study suggests that EFL teaching should pay attention to students' anxiety and take effective intervention measures, such as psychological counselling, emotion regulation skills training, etc., to help students reduce anxiety and improve learning performance.

5 Conclusions

This study focuses on EFL learners' anxiety factors and their impact on academic performance. A quantitative analysis method based on R language is proposed. Factor analysis and regression analysis models are used to explore the potential structure and mechanism of anxiety factors. Through factor analysis, it is identified that language anxiety, communication anxiety and test anxiety are the main factors affecting EFL learning, and language anxiety has the most significant impact on academic performance. This shows that in the process of foreign language learning, learners' anxiety in language expression is the core factor limiting their academic performance. Regression analysis showed that anxiety factors were significantly negatively correlated with academic performance. Specifically, learners with high levels of anxiety tend to show lower academic performance, while the impact of communication anxiety and test anxiety on academic performance is second, but it is still statistically significant. This shows that there is a close relationship between emotional state and academic performance. The results provide useful guidance for EFL teaching practice. Teachers should pay attention to students' emotional experience in the process of language learning, and design targeted strategies to reduce learners' anxiety level, so as to improve the teaching effect and students' academic performance.

Although this study reveals the basic structure and influence of anxiety factors in EFL learning, there are still some limitations.

The limitations of sample sources, the data are mainly from college students in a specific region, lack of extensive coverage of other learning groups, and the representativeness of the sample needs to be improved.

This study is based on cross-sectional data, which does not fully reflect the dynamic changes of learners' anxiety factors over time. Longitudinal research can make up for this deficiency in the future.

Based on the influence mechanism of anxiety factors, design intelligent teaching platforms or mobile learning applications to provide personalised intervention suggestions for different types of learners, helping them alleviate anxiety and improve learning efficiency.

Declarations

All authors declare that they have no conflicts of interest.

References

Bados, A., Gómez-Benito, J. and Balaguer, G. (2010) 'The state-trait anxiety inventory, trait version: does it really measure anxiety?', *Journal of Personality Assessment*, Vol. 92, No. 6, pp.560–567.

Bandelow, B., Michaelis, S. and Wedekind, D. (2017) 'Treatment of anxiety disorders', *Dialogues in Clinical Neuroscience*, Vol. 19, No. 2, pp.93–107.

Diamond, L.M. and Alley, J. (2022) 'Rethinking minority stress: a social safety perspective on the health effects of stigma in sexually-diverse and gender-diverse populations', *Neuroscience & Biobehavioral Reviews*, Vol. 138, p.104720.

Diep, L.T.N., Zainal, A.G., Hassan, I., Sunarti, Al-Sudani, A.Q.A.S. and Assefa, Y. (2022) 'Cultural familiarity, foreign language speaking skill, and foreign language anxiety: the case of Indonesian EFL learners', *Education Research International*, Vol. 2022, pp.1–9.

Fiddiyasari, A. and Pustika, R. (2021) 'Students' motivation in English online learning during Covid-19 Pandemic at SMA Muhammadiyah Gadingrejo', *Journal of English Language Teaching and Learning*, Vol. 2, No. 2, pp.57–61.

Fuller, R.M., Vician, C.M. and Brown, S.A. (2016) 'Longitudinal effects of computer-mediated communication anxiety on interaction in virtual teams', *IEEE Transactions on Professional Communication*, Vol. 59, No. 3, pp.166–185.

Hanifa, R. (2018) 'Factors generating anxiety when learning EFL speaking skills', *Studies in English Language and Education*, Vol. 5, No. 2, pp.230–239.

Hashemi, M. (2011) 'Language stress and anxiety among the English language learners', *Procedia-Social and Behavioral Sciences*, Vol. 30, pp.1811–1816.

Horwitz, E. (2001) 'Language anxiety and achievement', *Annual Review of Applied Linguistics*, Vol. 21, pp.112–126.

Horwitz, E.K. (1986) 'Preliminary evidence for the reliability and validity of a foreign language anxiety scale', *Tesol Quarterly*, Vol. 20, No. 3, p.559.

Jiang, Y. and Dewaele, J-M. (2019) 'How unique is the foreign language classroom enjoyment and anxiety of Chinese EFL learners?', *System*, Vol. 82, pp.13–25.

Li, X., Zhang, Y., Tiwari, P., Song, D., Hu, B., Yang, M., Zhao, Z., Kumar, N. and Marttinen, P. (2022) 'EEG based emotion recognition: a tutorial and review', ACM *Computing Surveys*, Vol. 55, No. 4, pp.1–57.

Liu, M. and Huang, W. (2011) 'An exploration of foreign language anxiety and English learning motivation', *Education Research International*, Vol. 2011, pp.1–8.

Liu, W. (2021) 'Does teacher immediacy affect students? A systematic review of the association between teacher verbal and non-verbal immediacy and student motivation', *Frontiers in Psychology*, Vol. 12.

Luo, H. (2013) 'Foreign language anxiety: past and future', *Chinese Journal of Applied Linguistics*, Vol. 36, No. 4, pp.442–464.

MacIntyre, P.D. and Gardner, R.C. (1991) 'Language anxiety: its relationship to other anxieties and to processing in native and second languages', *Language Learning*, Vol. 41, No. 4, pp.513–534.

Nielsen, B. and Harder, N. (2013) 'Causes of student anxiety during simulation: what the literature says', *Clinical Simulation in Nursing*, Vol. 9, No. 11, pp.e507–e512.

Pekrun, R., Marsh, H.W., Elliot, A.J., Stockinger, K., Perry, R.P., Vogl, E., Goetz, T., van Tilburg, W.A.P., Lüdtke, O. and Vispoel, W.P. (2023) 'A three-dimensional taxonomy of achievement emotions', *Journal of Personality and Social Psychology*, Vol. 124, No. 1, pp.145–178.

Russell, V. (2020) 'Language anxiety and the online learner', *Foreign Language Annals*, Vol. 53, No. 2, pp.338–352.

Spielberger, C. D., Anton, W. D. and Bedell, J. (2015) 'The nature and treatment of test anxiety', *Emotions and Anxiety: New Concepts, Methods, and Applications*, Vol. 10, No. 2, pp.317–344.

Svalberg, A.M.L. (2012) 'Language awareness in language learning and teaching: a research agenda', *Language teaching*, Vol. 45, No. 3, pp.376–388.

Szafranski, D.D., Barrera, T.L. and Norton, P.J. (2012) 'Test anxiety inventory: 30 years later', *Anxiety, Stress & Coping*, Vol. 25, No. 6, pp.667–677.

Tran, T.T.T. (2011) 'A review of Horwitz, Horwitz and cope's theory of foreign language anxiety and the challenges to the theory', *English Language Teaching*, Vol. 5, No. 1.

Vanden-Berghe, R., Verhagen, J., Oudgenoeg-Paz, O., Vander-Ven, S. and Leseman, P. (2018) 'Social robots for language learning: a review', *Review of Educational Research*, Vol. 89, No. 2, pp.259–295.

Vela, M.B., Erondu, A.I., Smith, N.A., Peek, M.E., Woodruff, J.N. and Chin, M.H. (2022) 'Eliminating explicit and implicit biases in health care: evidence and research needs', *Annual Review of Public Health*, Vol. 43, No. 1, pp.477–501.

Vonder-Embse, N., Jester, D., Roy, D. and Post, J. (2018) 'Test anxiety effects, predictors, and correlates: a 30-year meta-analytic review', *Journal of Affective Disorders*, Vol. 227, pp.483–493.

Yan, J.X. and Horwitz, E.K. (2008) 'Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: a qualitative analysis of EFL learners in China', *Language Learning*, Vol. 58, No. 1, pp.151–183.

Zwingmann, J., Baile, W.F., Schmier, J.W., Bernhard, J. and Keller, M. (2017) 'Effects of patient-centered communication on anxiety, negative affect, and trust in the physician in delivering a cancer diagnosis: A randomized, experimental study', *Cancer*, Vol. 123, No. 16, pp.3167–3175.