Foreign Language Classroom Anxiety Scale: A Comparison of Three Models

Rokiah Bt Paee • Jecky Misieng
Universiti Malaysia Sarawak

Abstract

Foreign language learning anxiety has long been recognized as one of the factors affecting the effectiveness of language learning but research findings have shown conflicting results. Horwitz et al. (1986) proposed a theory that predicts learners’ foreign language anxiety in the classroom and developed the Foreign Language Classroom Anxiety Scale (FLCAS) that was hypothesized to include three domains: communication apprehension, test anxiety and fear of negative evaluation. However Aida’s (1994) study revealed FLCAS is a four factor model: speech anxiety and fear of negative evaluation, fear of failing the Japanese class, degree of comfort when speaking with native speakers of Japanese and negative attitudes towards the Japanese class. Another four factor model was proposed by Zhao (2007) who reconstructed Horwitz et al.’s three factor model into four domains: communication apprehension, test anxiety, fear of negative evaluation and anxiety of foreign language class. In the current study, Horwitz’s three factor model, Aida’s and Zhao’s four factor model of the FLCAS were revisited and compared to see which one has a better fit for the Malaysian Japanese language learners. The FLCAS was administered to 328 beginning learners of Japanese in a Malaysian university. The internal consistency coefficient of the instrument was Cronbach’s alpha = .896 (m = 98.0 and SD = 15.17) which shows a reasonably high internal consistency. Confirmatory Factor Analysis (CFA) were carried using SPSS AMOS and Aida’s four factor model shows a better fit to the Malaysian data.

Keywo

Keyword: Foreign language anxiety, three factor model, four factor model, confirmatory factor analysis
1. Introduction

Foreign language learning anxiety has long been recognized and identified by instructors as one of the factors affecting the effectiveness of language learning. Defined as the dread or uneasiness that appears in a learner when attempting to use a foreign language by Gardner and MacIntyre (1993), it can also be viewed as a negative emotional response such as worry or anxious while in the process of learning and using a foreign language (MacIntyre, 1999). A large number of researchers from the field of language education and psychology have investigated the effects of anxiety on language education (Horwitz et. al., 1986; MacIntyre & Gardner, 1989; Skehan, 1989; Young, 1991) and their findings proved that some form of anxiety exist that inhibit learning of language especially a foreign language.

To date, many studies have shown a negative correlation between language anxiety and learner performance (Aida, 1994; Cheng et. al, 1999; Horwitz, 2001; MacIntyre & Gardner 1989; Saito & Samimy, 1996; Spielman & Radnofsky, 2001) with recent findings showing that anxiety can differentially affect how learners are able to process input and subsequently produce output (Sheen, 2008). Even though some studies have found that anxiety can have a facilitating effect on learning outcomes (Spielman & Radnofsky, 2001), most of these studies reported unaccommodating effect (Cheng at. al., 1999; Horwitz, 2001).

In explaining the varied conclusions reported by many researchers, Horwitz et. al. (1986) attributed them to the insufficient conceptualizations of the construct of anxiety made worse by the fact that there is a lack of valid and reliable anxiety measure specific to FL learning. Consequently, she and her colleagues developed an instrument called FLCAS to measure the level of learners’ anxiety. Horwitz et. al. (1986) then came up with the definition of language learning anxiety as an amalgamation of various incapacitating psychological as well as behavioral factors that go with language learning situations influenced by the unique process which is inherent in language learning. Based on this observation, three types of performance anxieties are proposed and labeled as communication apprehension, test anxiety and fear of negative evaluation. It is believed that these three anxieties form an integral part of FLA and together they conspire to inhibit learning as the learner attempts to learn and use a language.

It would seem that researchers have their finger on the crux of the matter pertaining to language learning anxiety with the advent of the FLCAS. However, a number of studies began to appear revealing inconsistent results with the sub categorization of language anxiety as well as the number of factors that made up the FLCAS. Exploratory studies that were carried out using the FLCAS have put forth findings that are both varied and inconsistent in the way the FLCAS measure language anxiety (Aida, 1994; Yuan, 2011; Zhao, 2007). Each study proposed
different models to capture the fundamental nature of language anxiety as proposed by Horwitz and her colleagues. One study (Gregersen, 2006) has shown that cultural differences may play a role in these mixed findings due to the fact that the FLCAS were used across language and cultural diversities. Therefore, this current study was an investigation to ascertain which of these proposed models can adequately explain FLA as defined by Horwitz et. al. (1986) based on data collected from students learning a foreign language in their native country.

2. Literature Review

The growing awareness experienced by many language educators on the debilitating effect of language anxiety has caused a significant rise in the number of studies regarding this phenomenon (Gregersen, 2006) both by language educators and psychologists. With the introduction of instruments such as the French Class Anxiety Scale and the French Use Anxiety Scale (Gardner, 1985) and the more popular FLCAS as developed by Horwitz et. al. (1986), interest in language anxiety research has continued to rise (Kitano, 1998).

Horwitz et. al. (1986) characterized language learning anxiety as ‘a distinct complex of self-perceptions, beliefs, feelings, and behaviors related to classroom language learning arising from the uniqueness of the language learning process’. In their conception, language learning anxiety, can be subcategorized into three distinct forms of performance anxieties namely communication anxiety, test anxiety and fear of negative evaluation. They argued that these three forms of anxiety are fundamental to the concept of FLA and together caused language learning to be a daunting task for a learner.

With this view of language anxiety, Horwitz and her colleagues came up with an instrument called the FLCAS. It was constructed based on self-reports from students, their own clinical experiences as well as evidence culled from reviews of similar instruments. The finalized version of the FLCAS contained 33 items which employs 5-point Likert-type scales with selections ranging from “strongly agree” to “strongly disagree.” This instrument was intended to measure foreign language learners’ level of anxiety while learning a language in the classroom. A higher score obtained by the instrument would indicate a higher level of FLA. Reliability of the scores obtained from the instrument based on data collected from 108 respondents was quite high with Cronbach’s alpha of .93 (Horwitz, 1986). In addition, test-retest reliability carried out with a sample of 78 participants over a period of eight weeks was ascertained to be $r= .83$ ($p < .01$) (Horwitz, 1991).
Aida (1994) decided to review Horwitz et al.’s three factor model of FLA by validating an adapted FLCAS based on students of Japanese to attempt to find out the underlying structure of the FLCAS. Assessment of reliability and the relationship of students’ anxiety levels to their performance in Japanese were carried out concurrently. The study reported that the adapted FLCAS was able to produce scores that are highly reliable at measuring anxiety levels of students learning Japanese. Aida established four factors instead of three: speech anxiety and fear of negative evaluation, fear of failing the Japanese class, degree of comfort when speaking with native speakers of Japanese, and negative attitudes towards the Japanese class. Six items (items 2, 6, 15, 19, 28 and 30) were removed from the final model because they failed to load on any of the four factors. Incidentally, the finding in this study was consistent with other studies based on western languages where they all concluded that language anxiety was negatively related to students’ performance in the language they were learning.
Another effort to re-examine Horwitz’s FLCAS factor structure was done by Zhao (2007) who used a Chinese version and established that there are also four factors namely communication anxiety, test anxiety, fear of negative evaluation and anxiety of English classes. The instrument was administered to a group of 115 second year high school students learning English as a foreign language in China. Unlike Aida’s four factor model, all 33 items from the original instrument were used in the final model proposed. However, Zhao did not report any reliability measures nor how the items load on their respective factors.

Figure 3. Model 3 FLCAS (Zhao, 2007)

Another researcher (Yuan, 2011) compared Zhao’s (2007) four factor model to that of Horwitz and colleague’s three factor model based on a sample of 300 Chinese college students learning English as a foreign language making use of the Chinese version of the FLCAS. In Yuan’s study, both models were reported to show all 33 items were significant indicators for their respective factors (p>.0001). The correlations between the four latent constructs for Horwitz’s model ranged from .91 to 1.0 while Zhao’s model has correlations ranging from .97 to 1.0. However, when the models were analyzed for fit, Horwitz and colleague’s model ($X^2=2169.18$, RMSEA=.07, AIC=7.72, SBC=8.57, BCCVI=7.78) was reported to have a much better fit for Chinese learners as compared to the Zhao’s model ($X^2=2211.19$, RMSEA=.07, AIC=7.88, SBC=8.77, BCCVI=7.93).

In Saito and Samimy’s study (1996), it was observed that language anxiety affected students differently in accordance to the level of instruction they were experiencing. They examined the role of language learner anxiety of Japanese language students’ as a predictor of performance at three different instructional levels. Among other predictor variables used was Year in College which was reported to be the best predictor of performance for beginning level
students. However, language anxiety came up on top as the best predictor of performance for both intermediate and advanced level students. The study proved conclusively that the predictive variable of the students of Japanese performance differs according to instructional levels and that the impact of FLA has a positive correlation with Japanese language learners' instructional levels.

Wan Zarina et. al. (2007) conducted a study on two groups of students in a Malaysian public university learning the Japanese language to determine and compare their perceptions towards their feelings of anxiety they experienced when learning the foreign language. Numerous variables that may affect the level of anxiety such as the student’s course, gender and race were studied. The findings of the study indicate that male students experienced higher level anxiety in communication apprehension compared to female students whereas in fear of negative evaluation, both gender experience the same level of anxiety. For the overall anxiety, Chinese and Malay students showed a higher level of anxiety compared to Indian students and students of other races.

In other Asian contexts, similar results were also obtained in studies involving both high school students and college students. In Taiwan, Cheng (1994) found out that scores obtained from the FLCAS was the best predictor of senior high school students’ English language proficiency levels. Meanwhile, in Korea, another researcher found FLCAS scores were significantly negatively correlated with final grades obtained by college English as Foreign Language learners in both reading and conversation classes (Kim, 1998)

3. Purpose of the Study

This study revisited Horwitz’s et. al. (1986) three factor model and Aida’s (1994) and Zhao’s (2007) four factor model of the FLCAS and compared it to see which model has a better fit for the Malaysian Japanese language learner.

4. Methodology

4.1 Participants

The participants of the present study were 328 beginner learners of Japanese at one public university in Malaysia. Japanese language course is offered as an elective subject and the students have to sit for four oral tests and one written test in order to pass the course. They were from various faculties, academic majors and ethnic groups.
4.2 Procedures

The questionnaire was distributed during the last week of the semester. At this point, all students had undergone the four oral tests. Therefore, learners’ FLCAS scores reflect their anxiety during the whole semester.

4.3 Instrument

Aida(1994) adapted version of Horwitz et. al. FLCAS where the term ‘foreign language’ in the original FLCAS were replaced with ‘Japanese language’ was used. The questionnaire consisted of two parts. Part one was intended to collect personal information of the participants, such as their name, age, gender, academic major and ethnicity and the participant’s first language. The second part consists of the 33 statements. The respondents were asked to rate each item on a five-point Likert scale ranging from 1 (strongly disagree) to 5 (strongly agree). When statements of the FLCAS were negatively worded, response were reversed and recoded.

4.4 Data Analyses

The reliability was tested using Cronbach’s Alpha and later confirmatory factor analysis (CFA) was done using SPSS AMOS version 20. Brown (2006) pointed out that in CFA, the number of factors and the pattern of indicator-factor loadings are specified in advance on the basis of strong empirical knowledge or theory. For the purpose of this study, three models form the basis the empirical knowledge and a comparison was carried out in terms of goodness-of-fit (GOF) statistics and size of the factor loadings to see which of the competing models have the best fit. The better fit model was determined by comparing the fit indices of the models produced using the value of Chi-square ($\chi^2$), Degree of Freedom ($df$), Root Mean Square Error Approximation (RMSEA), Akaike Information Criterion (AIC), Standardized Root Mean Residual (SRMR), Comparative Fit Index (CFI) and Parsimony Comparative Fit Index (PCFI).

For Chi-square goodness of fit index, the best fit is usually represented by small numbers, large p-values so as not to reject the null hypothesis of no significant difference between our observed and specified model. Other categories of fit indices are often recommended by experts to accompany the $X^2$ index and $df$ to ensure accuracy in our interpretation. They are either absolute or incremental fit indices with 1.0 as perfect fit and values above .95 as acceptable fit. Examples of absolute and incremental fit indices are CFI and PCFI Indices that are based on the discrepancy between observed and predicted covariance matrices of our models like the RMSEA and SRMR states that the value of 0 means perfect fit and values below .05 is considered adequate fit. The AIC, on the other hand, states that the model that gives small values close to 0 is more parsimonious and therefore, a better fitting model.
5. Results

The internal consistency coefficient of the instrument was Cronbach’s alpha = .896 ($m = 98.0$ and $SD = 15.17$) which indicates a reasonably high internal consistency.

Table 1

<table>
<thead>
<tr>
<th>A Comparison of the Three Models</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
</tr>
<tr>
<td>Aida (1994)</td>
</tr>
<tr>
<td>Horwitz et. al. (1986)</td>
</tr>
<tr>
<td>Zhao (2007)</td>
</tr>
</tbody>
</table>

Table 1 shows the fit statistics for all the three models. Aida’s model which shows both the CFI and PCFI indices as having higher values closest to the recommended .95 when compared to the Horwitz et. al.’s and Zhao’s models is the model of choice here. In terms of discrepancy between the specified and the observed models, all models show poor fit but according to the SRMR fit index, Aida’s model is not as bad as the other two models. However, based on the RMSEA, it is the worst fitting model. Based on $X^2$ values and $df$ to account for errors, Aida’s model is again on top of the other two models with less values and more degrees of freedom. In fact with a normed Chi-square (NC) ratio where the $X^2$ is divided by its $df$) of 2.49 indicative of a parsimonious good fit, it is far better than Horwitz et. al.’s and Zhao’s models with NC’s of 2.38 and 2.40 respectively. This is further supported by the smallest AIC value that the model obtained as compared to the other two models. Therefore, all fit indices except the RMSEA reveals that Aida’s model is the best fitting model.

6. Discussion and Conclusion

FLA has become the focus of many language learning research and findings have shown conflicting result. Horwitz et. al. (1986) hypothesized a theory that predicts learners’ FLA in the classroom setting and came up with a 33 item instrument called the FLCAS to measure it. Based on the instrument, they proposed a three factor model that includes the three domains of communication apprehension, test anxiety and fear of negative evaluation. Using an Asian language to tests Horwitz et. al.’s conception of FLA. Aida conceptualized a four factor model but did not include six items due to weak factor loadings on their respective factors. The domains of Aida’s (1994) model include speech anxiety and fear of negative evaluation, fear of failing the class, comfortableness in speaking with native Japanese and negative attitudes towards the Japanese class. Zhao (2007), on the other hand, retained all the original 33 items in his four factor model which he renamed as communication anxiety, test anxiety, fear of negative evaluation and anxiety of English classes. All the researchers reported high
reliabilities in the scores obtained from their instruments.

Based on the confirmatory factor analyses carried out on all the three models, Aida’s four factors model is discovered to be a more viable model in accounting for FLA among Malaysian students studying Japanese as a foreign language. The best possible explanation for this is that unlike the other two models, in Aida’s model, six items that were not loading particularly well on their factors were removed from the analysis. Upon closer examination, many of the items were either too similar or redundant. Examples of the items included are item 19, “I am afraid that my Japanese teacher is ready to correct every mistake I make” and item 15, “I get upset when I don’t understand what the teacher is correcting”. Furthermore, the factor of Test Anxiety which was present in both Horwitz et. al.’s and Zhao’s model but not found in Aida’s model could be another possible explanation for this result. This seem to confirm the finding of MacIntyre & Gardner (1989) that test anxiety contributed to the general anxiety factor and not to the communicative anxiety factor which seems to indicate that test anxiety is a more generic predicament faced by students instead of a specific indicator of FLA.

A follow up study is recommended to test more models that have been proposed to seek the best model of all that can best explained this complex phenomenon.

References

50


