

# 外国語教室不安尺度 (FLCAS) : 3 つのモデル比較再考

ダイクス ロバート<sup>\*</sup>・ピロット クリストファー<sup>\*\*</sup>

仁愛大学人間学部<sup>\*</sup>・福井工業大学 基盤教育機構<sup>\*\*</sup>

## FLCAS: A Comparison of Three Models Revisited

Robert DYKES<sup>\*</sup> & Christopher PIROTTO<sup>\*\*</sup>

Faculty of Human Studies, Jin-ai University<sup>\*</sup>・Organization for Fundamental Education, Fukui University of Technology<sup>\*\*</sup>

本論文は Paee and Misieng (2012) が比較検討した 3 つの外国語教室不安尺度 (FLCAS) モデルのレプリケーションである。本論文では Paee and Misieng (2012) の行った比較を、英語を学習する日本人大学生にとって最適なモデルという観点から再比較を行っている。確証的因子分析によって得られた本論の検討結果は Paee and Misieng (2012) と同意するものだったが、それと同時に FLCAS のような便利な予め用意された測定方法は、考察・分析された後、使用される環境によっては調整の必要があるという問題点も明らかにした。つまり、FLCAS を学習者に対して測定手段として用いる場合、研究のコンテキストに最も適した因子スケールを決定するためには、最低でも探索的因子分析を行うべきであろう。本論文はまた、FLCAS への変更は統計的分析に基づいて行われ、それらの変更は適宜に記録されなければならないという点も強調している。

キーワード：FLCAS, 外国語教室不安尺度, 第二言語不安, フィット統計

This paper is a replication of the Paee and Misieng (2012) comparison of three popular Foreign Language Classroom Anxiety Scale (FLCAS) models. This study applies the Paee and Misieng (2012) comparison to a different context of students to determine which is the best fitting FLCAS model for Japanese university students learning English. The findings of this paper, from a confirmatory factor analysis, concur with Paee and Misieng in addition to bringing forth the issue that readymade measurement tools such as the FLCAS, while convenient, need to be examined, tested, and often tweaked for various usage contexts. It is recommended that researchers that wish to utilize the FLCAS as a measurement tool for their students should at a minimum conduct an exploratory factor analysis to determine the best factor scale for their study's context. This paper also highlights the importance that changes to the FLCAS should be supported by statistical analysis and documented accordingly.

Keywords: FLCAS, Foreign Language Anxiety, fit statistics

## Introduction

The Foreign Language Classroom Anxiety Scale (FLCAS; Horwitz, Horwitz, & Cope, 1986) has been used extensively to measure foreign language anxiety (FLA) in classrooms around the globe. FLA research has revealed some conflicting results (causes, test outcomes, and remedies to reduce FLA) and it has been argued that some of this contention is the result of the design of the FLCAS or in some cases the model used to examine the results from the FLCAS. Horwitz et al. (1986) proposed a 3-factor model comprised of communication apprehension, test anxiety, and fear of negative evaluation (see Figure 1). Aida (1994) reexamined the FLCAS and developed a 4-factor model composed of speech anxiety and fear of negative evaluation, fear of failing the language class, degree of comfort when speaking to native L2 speakers, and attitudes toward L2 class (see Figure 2). An additional 4-factor model was also proposed by Na (2007) which is comprised of communication apprehension, test anxiety, fear of negative evaluation, and anxiety of L2 class (see Figure 3). Paee and Misieng (2012) examined these 3 models through confirmatory factor analysis (CFA) and determined that the Aida (1994) model was the best fit for their FLA study involving learners of Japanese in a Malaysian university.

This paper revisits the Paee and Misieng (2012) comparison and applies their methodology to the results of a previous study which measured the FLA of Japanese university students learning English (see Dykes, 2018). This paper reached the same conclusion as Paee and Misieng (2012) that the Aida (1994) model is the best fitting model, of these three models, to apply to FLCAS results. However, after further review of FLCAS literature, an additional conclusion was reached that while the Aida (1994) model was the best fitting model in 3 studies, Aida (1994), Paee and Misieng (2012), and this study, there is no such thing as a best fitting FLCAS model for FLA studies. This

led to the final conclusion that because of the cultural and contextual sensitivity of the FLCAS, a factor analysis should be run for each and every project utilizing the FLCAS and a situational model created for each and every project. (It should be noted that Na (2007) was incorrectly cited in Paee and Misieng (2012) as Zhao (2007).)

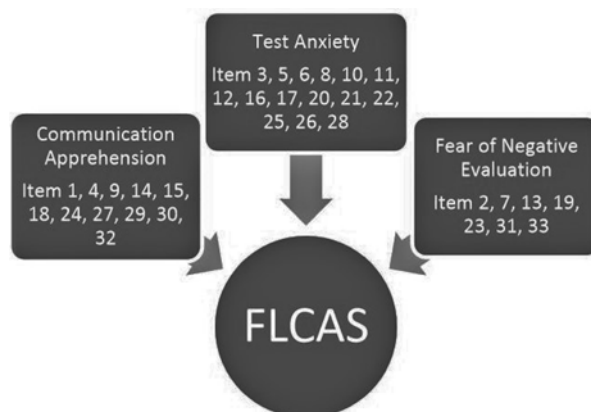


Figure 1: Horwitz (1986) model adapted from Paee and Misieng (2012)

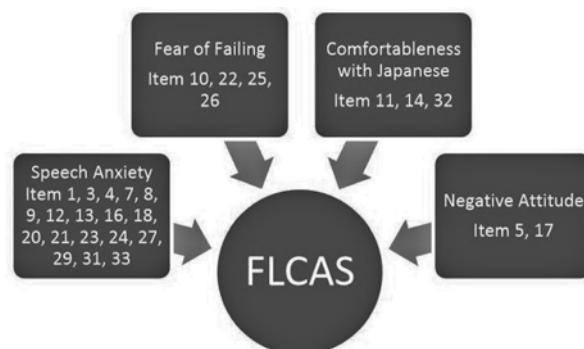


Figure 2: Aida (1994) model adapted from Paee and Misieng (2012)

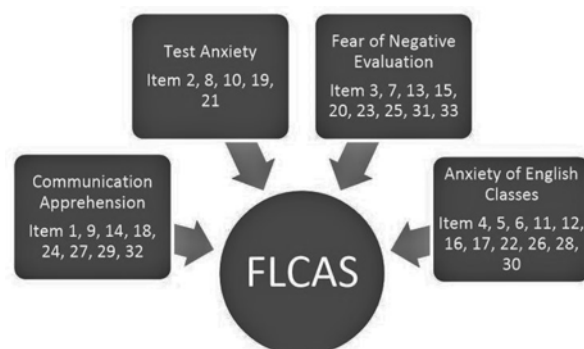


Figure 3: Na (2007) model adapted from Paee and Misieng (2012)

## Literature Review

This literature review will first briefly outline the history leading up to development of the FLCAS. It will then examine the FLCAS's use in Japan which will also include issues in its usage and what some researchers have done to combat these issues. For a more thorough synopsis of the Horwitz et al. (1986), Aida (1994), and Na (2007) papers, see Paee and Misieng (2012).

Language anxiety research is usually traced back to Dunkel (1947) who noticed that students' anxious thinking negatively correlated with proficiency in Latin class. It then saw a more formal entry into applied linguistics with the publication of Chastain (1975), who examined affective factors in second language learning including anxiety, and Scovel (1978) who performed a review on anxiety in the foreign language learning context. However, we also begin to see claims that inadequate, inconclusive results, and unreliable conclusions are being published on FLA. A review of the literature today shows that such issues still exist which is why Paee and Misieng and this present paper are still trying to examine a 35-year-old measurement tool. That tool is the FLCAS which was born out of the same desire to be able to better measure the effects of FLA. In the paper that introduced the FLCAS, Horwitz et al. (1986) state that part of the motivation to create the FLCAS was because efforts to measure and quantify FLA prior to the FLCAS, "have met with mixed results" (p. 125). Up to this point the only measurement tool similar to the current FLCAS was a subtest of the Attitude/Motivation Test Battery labeled "French Class Anxiety" which was used in a number of studies beginning in 1975 (Gardner, 1985; MacIntyre and Gardner, 1991). In MacIntyre and Gardner (1991) an extensive literature review was conducted on the FLA field examining some major models and frameworks (e.g., The French Class Anxiety Scale, English Use Anxiety, English Test Anxiety, and the FLCAS). They

note that up to this point it had been firmly established that FLA can impair language learning and production. MacIntyre and Gardner (1991) write rather favorably of Horwitz et al. and their paper on the FLCAS stating "the literature on foreign language anxiety generally supports the tenets advanced by Horwitz et al. (1986)" (p. 105). MacIntyre and Gardner (1991) do, however, identify an issue with Horwitz et al.'s framework which later received much more critical attention in Aida (1994); the second component of the FLCAS, test anxiety. MacIntyre and Gardner (1991) note that "Horwitz et al. (1986) are not clear whether this test anxiety is specific to the type of tests encountered in language class or whether it is a generalized test anxiety" and conclude that "[t]est anxiety may be less important, although a rigorous investigation has yet to be made" (pp. 105-106). This is an important question to explore because if test anxiety is a generalized form of anxiety and not specific to FLA or FLA testing, then its inclusion in the FLCAS needs to be removed and the Likert items adjusted. This concern is exactly what Aida (1994) undertakes and with it conducts a critical examination of the original FLCAS model and concludes that test anxiety is indeed not a factor of FLA but a component of generalized test anxiety as MacIntyre and Gardner (1991) had hinted at.

In Kawashima (2009) a thorough review was conducted on FLA articles published in the context of Japan. Kawashima found that the FLCAS (modified or unmodified) was by far the most popular instrument used to measure FLA. Out of twenty-six studies, 10 used the FLCAS, and the second most used instrument, the Foreign Language Reading Anxiety Scale (FLRAS), was only used in 5 studies. Just a brief search on FLA research conducted in Japan since Kawashima's article returns nearly a dozen new research articles using the FLCAS, leading one to believe that despite its age, the FLCAS remains a very popular tool in FLA research. The reliability and validity of the Horwitz et al. FLCAS has been brought

into question for its use in Japan, Asia, and even its native North America. However, the FLCAS is often used in its original form or not modified following proper methods that are replicable (e.g., not offering any reason for changes, not stating clearly what was modified or changed, or not publishing the data to back up any claims for the reasons of modification). In Japan, three cases of FLCAS usage are prevalent. (A) Adequate modification that is explained, documented, and replicable, (B) no modification of the original model, and (C) modification with no justification given or inadequate reporting so that replication and/or verification is impossible. Kawashima (2009) points out some papers modified the FLCAS but did not adequately document the changes and therefore render the studies nonreplicable:

Matsumura (2000) used a 7-point scale in her study, but did not provide a specific rationale for this... [and] the reasons for deleting certain items were not always clearly reported. McLaughlin and Yamashiro (2000) discarded two items because they substantially reduced reliability of the scale and they seemed different from other items in the construct [McLaughlin and Yamashiro (2000) took the correct steps for modification but did not publish the results of their factor analysis so their claims of reliability could not be verified or reproduced]. Asano (2003) selected 23 items from the original 33 without giving a rationale. Matsumura (2000) added 17 items to the FLCAS based on the results of an open-ended questionnaire and informal interviews she had conducted previously. (p. 237)

Regarding the FLCAS subscales Kawashima (2009) continues with “[d]ue to incomplete reporting of data, reliability figures for the scales cannot be compared further” (p. 238). One such example includes a FLCAS item scale that was classified differently by Falout (2004) and McLaughlin and Yamashiro (2000) and the

only justification for their respective classifications appearing to be personal interpretation making internal consistency incomparable.

Kawashima (2009) remarks that some important modifications can be noted such as changing the items that refer to “speaking”. For example, Asano (2003) made this type of modification based upon observations of the teaching environment they would be testing. “[W]hen they speak English” was changed to “when they read aloud and speak English” (Asano, 2003). Kawashima reports that despite Kondo and Yang (2003), Matsuda and Gobel (2004), and Yashima (2002) all claiming that the FLCAS in its unaltered form is inappropriate for use in Japan without modification, “few scales specifically modified for the Japanese context have been documented” (Kawashima, 2009; p. 239). The main reason the FLCAS is inappropriate for use in Japan (or anywhere, for that matter), in an unaltered form is due to its sensitivity to both cultural and contextual changes (e.g., teaching environment, language ability level, teaching method, learning method, etc.) (Toyama and Yamazaki, 2018).

Following the Kawashima (2009) review there have been more studies using the FLCAS in the Japanese context. Examining these studies, it seems some still did not modify their FLCAS or adequately document the changes made. However, it appears that more researchers are modifying the FLCAS (and documenting the changes) than the older studies covered in the Kawashima (2009) review. Some notable examples include Horai (2012) which offers up a Japanese translation of the FLCAS including a thorough analysis, which ensures the suitability of the translation. Horai (2012) provides both a Kaiser-Meyer-Olkin test as well as an exploratory factor analysis (EFA) of the FLCAS (an EFA can assure that the item scales remain unaltered by possible changes through translation). Iimura (2016a; 2016b) uses the Yashima et al. (2009) translated version of the FLCAS and reports no modifications or changes were made.

Saito and Iida (2016) translated and back-translated the FLCAS and note they “modified some parts to adjust it to the Japanese participants” (p. 10), but do not note what these changes are and do not include the translation in the published article. They do, however, run an EFA providing adequate data to support their three-factor scale. Mathieson (2016) does modify the FLCAS for his context, but unfortunately it appears to be based on personal interpretation, rather than a statistical analysis such as EFA. Dykes (2017) opted for the Yashima et al. (2009) version of the FLCAS and changed the 5-point Likert scale to a 6-point scale, as well as including the justification for such a modification. Pirotto (2018) translated and back-translated the FLCAS and included a copy of the translated version so that any changes can be examined. An EFA was also conducted with adequate data to back up factor scaling and deletion of items.

### **Purpose of the Study**

This study’s main goal was to check the reproducibility of Paee and Misieng’s (2012) paper which consists of CFA performed on three popular FLCAS models. The three models examined are those by Horwitz et al. (1986), Aida (1994), and Na (2007). Paee and Misieng (2012) concluded that Aida’s model was the best fitting model for their teaching context. Therefore, this research paper will determine if Aida’s model is still the best fit for a different, completely unrelated, group of students. In order to determine model fit, CFA was run through AMOS software and several fit indices were then compared.

### **Methodology**

#### **Participants**

The participants for this replication study were Japanese university learners of English (n=593). Their estimated level of English (determined by TOEIC scores, custom CEFR interviews, and entrance exam results) is around CEFR A1 to CEFR A2. They

were all enrolled in a 16-week mandatory English University course that met twice a week for 90 minutes per class for a total of 28 classes (some weeks the students only met for class once). Participants represented 15 different academic departments and majors and were taught by a total of 12 different English language instructors. For a more detailed breakdown of the participants see Dykes (2018).

#### **Instrument**

The instrument (FLCAS) used in this study, Aida (1994), Na (2007), and Paee and Misieng (2012) was derived from the original 33-item FLCAS originating from Horwitz et al. (1986). This present study employed a Japanese translation version taken from Yashima et al. (2009). The terminology “foreign language class” was left as is in this version. It was, however, changed to “Japanese class” in Aida (1994) and Paee and Misieng (2012). This version of the FLCAS includes all 33 items, whereas Aida (1994) removed six items. It should also be noted that for the present study the 5-point Likert scale was changed to a forced choice hybrid 6-point scale. This was done in order to eliminate neutral responses. The neutral response is used to indicate ambivalence, indifference or confliction of an item or question, which is often appropriate for many forms of research, such as a marketing study, but not for the scale items on the FLCAS. A secondary reason for this choice is the Asian, Japanese context in which the FLCAS was administered. Wang, Brain, Dugan, and Komives (2013) theorizes Japanese respondents often chose the middle option, (when an odd number of choices is given) due to a cultural component that is prevalent in collectivist societies like Japan. An additional study which examined the survey response habits of 26 different countries found that Japan had the highest prevalence to choose the middle answer, (when an odd number of choices is given) (Harzing, 2006). The 6-point scale ranges from strongly agree = 1 to

strongly disagree = 6. Items 2, 5, 8, 11, 14, 18, 22, 28, and 32, on the FLCAS were negatively worded and their scores were reversed and recoded for the data analysis portion of this project.

### Data Analysis

CFA was conducted using SPSS AMOS (v. 23). Some researches may not agree with specific fit indices used in this paper or even the quantity included, as some may feel they are redundant or unnecessary. However, this is a replication study so the same fit indices that Paee and Misieng (2012) used are also included in this study to determine the better fit model. The values included are Chi-square ( $X^2$ ), 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). An additional fit index was also included as recommended by Kenny (2015), the Tucker Lewis Index (TLI). The TLI for all studies was able to be calculated based on the data already available.

For the Chi-square goodness of fit test, smaller values typically represent a better fit. This is usually examined in conjunction with the degrees of freedom. RMSEA and SRMR are examples of fit indices that have a value of 0 being perfect fit and values below .05 being considered adequate fit. AIC also considered lower values to be better fitting. On the other hand, fit indices such as CFI, PCFI, and TLI have a value of 1.0 being perfect fit and anything above .95 being considered adequate fit.

### Results

The Aida (1994) model is clearly shown as the best fitting model based on several results seen in Table 1. Examining the  $X^2$  value, Aida's model is a better fit with lower values and more degrees of freedom (the  $df$  value is included to ensure accuracy of the interpretation of the data). Aida's model also has CFI, PCFI and TLI results higher than, and closer to 1, than the Horwitz et al. (1986) or Na (2007) models. It should be noted that while Aida's model indicates better fit, due to the CFI, PCFI, or TLI results not reaching 0.95, the results are not considered statistically significant. It can even be argued that none of these models reach an adequate level of fit. The last indicator of a better fit for Aida's model is the much lower AIC value when compared to the other two models. The AIC value is about 1000 points lower for Aida's model than both Horwitz et al. and Na's models. The only values where Aida's model does not indicate a better fit are the RMSEA and SRMR indices. When rounded to the 2nd decimal place all models had RMSEA values of 0.08 and SRMR values of 0.07. When using the RMSEA and SRMR indices no conclusion can be drawn as to which model is better fitting. It should be noted that while the RMSEA and SRMR values were very low for all three models no value was below 0.05, indicating a lack of statistical significance for all three models.

### Discussion and Conclusion

The motivation to conduct this paper and continue looking further at the FLCAS is driven by the authors' own work with the FLCAS in Japan and with sentiments such as those expressed by Kawashima (2009), "the validity and reliability of

Table 1: FLCAS CFA Model comparison

Model	$X^2$	$df$	RMSEA	SRMR	AIC	CFI	PCFI	TLI
Horwitz et. al. (1986)	2402.26	492	0.08	0.07	2540.26	0.78	0.73	0.76
Aida (1994)	1390	318	0.08	0.07	1510	0.84	0.77	0.83
Na (2007)	2367.05	489	0.08	0.07	2511.05	0.78	0.73	0.77

the FLCAS when administered in Japan is doubtful. High reliability and validity of the scale in studies conducted abroad does not necessarily guarantee high reliability and validity in Japan” (p. 235). The simple conclusion to be drawn from the Paee and Misieng (2012) study, and backed up by this replication study, is that the Aida (1994) model, when using certain fit indices, appears to be a better fitting model when compared to Horwitz et al.’s (1986) and Na’s (2007) models.

Paee and Misieng (2012) conclude their paper with “[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[sic] this complex phenomenon” (p. 50). This, however, would not be a fruitful endeavor as there is no best fitting model for all contexts, or even Japanese contexts. This becomes abundantly clear in Toyama and Yamazaki (2018) where eight FLCAS models are examined. They found no two models contained the same item scales. There was no agreement with how many factors an FLCAS model should include (2, 3, and 4-factor models) and no consistency of items within any given factor. The main conclusion to be drawn from their extensive study is the FLCAS is a very sensitive measurement tool that is affected by “cultural contexts, student’s major, translation, and analytical strategies” (Toyama and Yamazaki, 2018; p. 21). Other factors that seem to affect the FLCAS as seen in the examples from studies covered in the literature review include the classroom and method of instruction, most notably how much and what kind of speaking will take place. This can often differ greatly from a K-12 classroom to the more communicative classrooms often found in a Japanese university.

In conclusion, it is recommended that researchers that wish to utilize the FLCAS as a measurement tool for their students should at a minimum conduct a confirmatory factor analysis which can be used to determine which (if any) existing model fits their

learners the best. Instead of assuming one model fits every teaching context, it is highly recommended that several models are tested. If no existing models are a good fit, it is suggested that an exploratory factor analysis leads you to a different and better fitting model. EFA is used to determine the best factor scale for a study’s context. EFA will help to find groupings of highly correlated questionnaire items in order to create a factor model which can be used to interpret the data from the FLCAS.

It is also important to check the translation to determine if it fits the context for which it is being applied to. It may be necessary to create a new translation for the specific context that it is being applied to. This should be done with caution to ensure that new translations, retranslations, or edits to translations are specific, reliable, and accurate. This will not only increase the validity of the research but also will be helpful to researchers who may want to use that specific translation in the future. In order to ensure translations are reliable and accurate, it is recommended these steps are followed:

- 1) Translate into Japanese
- 2) Have different individuals back translate to English
- 3) Compare, adjust, and repeat until similar translations are achieved
- 4) Pilot the translation and test for clarity
- 5) Include your translation in your work and note any changes made from the original

Noting and including any changes or modifications is also crucial. One of the most important things about conducting research is to make research replicable. Therefore, it is encouraged that anyone who uses the FLCAS should clearly document all changes made to the FLCAS and justify why these changes were made. This includes the removing or adding of Likert items, custom translations, justification for models

used, how many points Likert items are worth, and the Likert scale used. If possible, statistical analysis of the FLCAS data should be included to justify the decisions made. Including all of this will ensure that future researchers are able to understand exactly how the study was done, ensuring replicability.

To further improve the FLCAS for application in Japan a follow-up study is recommended to examine the number of Likert choices used in the FLCAS. Fallout (2004) uses a 6-point scale and Matsumura (2000) uses a 7-point, but neither offer any justification. Dykes (2017) offers an argument for a 6-point scale to be used in Japanese context, but a thorough statistical analysis would provide more concrete answers to the issue.

## References

- Aida, Y. (1994). Examination of Horwitz, Horwitz, and Cope's construct of foreign language anxiety: The case of students of Japanese. *The modern language journal*, 78(2), 155-168. doi: 10.2307/329005
- Asano, S. (2003). Some effects of audio-visual approaches on foreign language learning anxiety. *Bulletin of Osaka University of Health and Sports Sciences*, 34, 65-82. Retrieved from <https://ci.nii.ac.jp/els/contents110004688994.pdf?id=ART0007424831>
- Dykes, R. (2017). The Effect of Communicative Based Teaching on Foreign Language Anxiety. *International Education and Exchange Research*, 1, 37-49. Retrieved from [http://news2.ad.u-fukui.ac.jp/wp/wp-content/uploads/4\\_DYKES.pdf](http://news2.ad.u-fukui.ac.jp/wp/wp-content/uploads/4_DYKES.pdf)
- Dykes, R. (2018). Syllabus Design and Foreign Language Anxiety. In P. Clements, A. Krause, & P. Bennett (Eds.), *Language teaching in a global age: Shaping the classroom, shaping the world*. Tokyo: JALT. Retrieved from <https://jalt-publications.org/sites/default/files/pdf-article/jalt2017-pcp-009.pdf>
- Chastain, K. (1975). Affective and ability factors in second language acquisition. *Language Learning*, 25(1), 153-161. <https://doi.org/10.1111/j.1467-1770.1975.tb00115.x>
- Dunkel, H. (1947). The effect of personality on language achievement. *Journal of Educational Psychology*, 38(3), 177-182. <https://doi.org/10.1037/h0057305>
- Falout, J. J. (2004). Foreign language anxiety: The case of supplementary English classes at the College of Science and Technology, *Nihon University. Bulletin of Department of General Education, College of Science and Technology, Nihon University*, 75, 1-8. Available from <https://ci.nii.ac.jp/naid/40006261942>
- Gardner, R. C. (1985). The Attitude/Motivation Test Battery: Technical report. Retrieved from <http://publish.uwo.ca/~gardner/docs/AMTBmanual.pdf>
- Harzing, A. (2006). Response Styles in Cross-national Survey Research A 26-country Study *International Journal of Cross Cultural Management*, 6(2), 243-266.
- Horai, K. (2012). A measurement study of a Japanese translation of the Foreign Language Classroom Anxiety Scale. *Bulletin of Sojo University*, 37, 109-119. Retrieved from <https://ci.nii.ac.jp/naid/110009422542>
- Horwitz E., Horwitz, M., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, 70(2), 125-132. <https://doi.org/10.2307/327317>
- Iimura, H. (2016a). Anxiety and the Effect of Frequent Contact With Target Language Speakers in Japanese EFL Context. *The Chubu English Language Education Society*, 45, 31-36. [https://doi.org/10.20713/celes.45.0\\_31](https://doi.org/10.20713/celes.45.0_31)
- Iimura, H. (2016b). The Effect of an Oral Presentation Contest on Students' Anxiety in Japanese EFL Context. *KATE Journal*, 30, 71-84. [https://doi.org/10.20806/katejournal.30.0\\_71](https://doi.org/10.20806/katejournal.30.0_71)
- Kawashima, T. (2009). FL anxiety studies of Japanese EFL learners: A critical review. In A. M. Stoke (Ed.), *Shared identities: Our interweaving threads*. Tokyo: JALT. Retrieved from <http://jalt-publications.org/recentpdf/proceedings/2008/E120.pdf>
- Kenny, D. (2015). Measuring Model Fit. Retrieved from <http://davidakenny.net/cm/fit.htm>
- Kondo, S., & Yang, Y. (2003). The English language classroom anxiety scale: Test construction, reliability, and validity. *JALT Journal*, 25(2), 187-196. Retrieved from <https://www.jalt-publications.org/files/pdf-article/jj-25.2-art4.pdf>
- MacIntyre, P., & Gardner, R. (1991). Methods and results in the study of foreign language anxiety: A review of the literature. *Language Learning*, 41, 85-117. <https://doi.org/10.1111/j.1467-1770.1991.tb00677.x>
- Mathieson, P. (2016). Foreign Language Anxiety in Community English Conversation Classes in Japan. *OSAKA JALT*, 88. Retrieved from [http://www.osakajalt.org/storage/Osaka\\_JALT\\_Journal\\_2016.pdf#page=89](http://www.osakajalt.org/storage/Osaka_JALT_Journal_2016.pdf#page=89)

- Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System*, 32, 21-36. <https://doi.org/10.1016/j.system.2003.08.002>
- Matsumura, Y. (2000). The effects of foreign language classroom anxiety on a listening test for Japanese EFL learners of English. *Journal of Linguistic and Cultural Studies, Kinki University*, 15, 41-63. Available from <https://ci.nii.ac.jp/naid/40005018678>
- McLaughlin, J. W., & Yamashiro, A. D. (2000). The relationships between English language proficiency and foreign language classroom anxiety among Japanese university students. *The Journal of Heisei International University*, 4, 95-114. Available from <https://ci.nii.ac.jp/naid/40005164383>
- Na, Z. (2007). A study of high school students' English learning anxiety. *The Asian EFL Journal*, 9(3), 22-34. Retrieved from [http://asian-efl-journal.com/September\\_2007\\_EBook\\_editions.pdf](http://asian-efl-journal.com/September_2007_EBook_editions.pdf)
- Pace, R. B., & Misieng, J. (2012). 外国語教室不安尺度—3つのモデルの比較. 第9回マレーシア日本語教育研究発表会, 42-51. Retrieved from <http://www3.grips.ac.jp/~jlc/jlc/oversea/rokiahreport.pdf>.
- Pirotto, C. (2018). A factorial investigation into the English language learning anxiety of first-year university students in Japan. *Memoirs of Fukui University of Technology*, 48, 165-172. Available from <https://ci.nii.ac.jp/naid/40021594396>
- Saito, T., & Iida, T. (2016). Speaking Anxiety and the Effects of Previous Overseas Experience in an English Intensive Program before Study Abroad. *Papers in Language, Literature, and Culture of the Graduate School of Doshisha Women's College of Liberal Arts*, 15, 1-24. doi/10.15020/00000879
- Scovel, T. (1978). The effect of affect on foreign language learning: A review of the anxiety research. *Language Learning*, 28(1), 129-142. <https://doi.org/10.1111/j.1467-1770.1978.tb00309.x>
- Toyama, M., & Yamazaki, Y. (2018). Exploring the components of the foreign language classroom anxiety scale in the context of Japanese undergraduates. *Asian-Pacific Journal of Second and Foreign Language Education*, 3(1), 4. <https://doi.org/10.1186/s40862-018-0045-3>
- Wang, R., Hempton, B., Dugan, J., & Komives, S. (2013). Cultural Differences: Why Do Asians Avoid Extreme Responses? *Survey Practice*, 1(3).
- Yashima, T. (2002). Willingness to communicate in a second language: The Japanese EFL context. *The Modern Language Journal*, 86(1), 54-66. <https://doi.org/10.1111/1540-4781.00136>
- Yashima, T., Noels, K., Shizuka, T., Takeuchi, O., Yamane, S., & Yoshizawa, K. (2009). The interplay of classroom anxiety, intrinsic motivation, and gender in the Japanese EFL context. *Journal of Foreign Language Education and Research*, 17, 41-64. Retrieved from <https://kuir.jm.kansai-u.ac.jp/dspace/bitstream/10112/768/1/KU-1100-2009300-.pdf>

## Appendix

### 外国語教室不安尺度 (FLCAS)

(本研究において \*は逆転項目として扱った.)

- 1) 外国語の授業で話すとき自信がもてない.
- 2) 外国語の授業で間違えることは気にならない. \*
- 3) 外国語の授業で当てられると思うと体が震える.
- 4) 外国語の授業で先生の言っていることが理解できないととても不安だ.
- 5) もっと外国語の授業があってもよいと思っている. \*
- 6) 外国語の時間授業と関係ないことを考えていることがよくある.
- 7) 他の生徒の方が自分よりよくできていると思っている.
- 8) 外国語の授業中のテストではだいたい落ち着いている. \*
- 9) 外国語の授業で準備なしに話さないといけない時、パニックになる.
- 10) 外国語の単位を落としたときの影響が心配だ.
- 11) 外国語の授業で動揺する人の気持ちがわからない. \*
- 12) 外国語の授業では、緊張のあまり、知ってたことも忘れてしまうときがある.
- 13) 外国語の授業で自分からすすんで答えるのは恥ずかしい.
- 14) 外国語をネイティブスピーカーと話すとき緊張しない. \*
- 15) 先生が何を訂正しているのか理解できないとき動揺する.
- 16) 外国語の授業の予習を十分にしていっても心配になる.
- 17) よく外国語の授業を休みたくなる.
- 18) 外国語の授業で話すのに自信がある. \*
- 19) 先生が自分の間違いをいちいち直そうなので心配だ.
- 20) 外国語のクラスで当たりそうになると胸がどきどきする.
- 21) 外国語のテスト勉強をすればするほど、混乱する.
- 22) 外国語の授業の予習をよくしないといけないというプレッシャーは感じない. \*
- 23) 常に他の学生の方が外国語で話すのが上手だと感じている.
- 24) 他の学生の前で外国語を話すとき自意識がとても高くなる.
- 25) 外国語のクラスは進むのが速いのでついていけないかどうか心配である.
- 26) 他の科目よりも外国語のクラスの方が緊張する.
- 27) 外国語のクラスで話すとき緊張したり混乱したりする.

- 28) 外国語のクラスに向かうとき自信をもてるしリラックスしている. \*
- 29) 先生の言うことがすべて理解できないと不安になる.
- 30) 外国語を話すためにあまりに多くの文法規則を勉強しないといけなので圧倒される.
- 31) 私が外国語を話すと他の学生が笑うのではないかと思う.
- 32) ネーティブスピーカーに会うときおそらくリラックスしていられると思う. \*
- 33) 先生が, 前もって準備していなかった質問をすると緊張する.

**FLCAS: Howritz (1986)**

**Translation Yashima et al., (2009)**