

REPLICATION RESEARCH

Approximate replication of Matsuda and Gobel (2004) for psychometric validation of the Foreign Language Reading Anxiety Scale

Akira Hamada^{1*} and Shuichi Takaki²

¹Meikai University, Japan and ²Fukushima University, Japan

*Corresponding author. E-mail: hamada.akira@meikai.ac.jp

Abstract

This paper reports an approximate replication of Matsuda and Gobel (2004) for the psychometric validation of the Foreign Language Reading Anxiety Scale (FLRAS). Their study examined the structural aspects of the FLRAS developed by Saito, Horwitz, and Garza (1999). The results showed that the FLRAS measured three different subcomponents of foreign language reading anxiety; none of the factors predicted foreign language performance in content-based and four-skill classes. The present study aimed to reconfirm the psychometric validity of the FLRAS because it has been widely employed to make foreign language reading anxiety researchable. Our study retained the same methodology, with the exception of the measurements of classroom performance and reading proficiency. Matsuda and Gobel's conclusions were reproduced by showing a weak relationship between classroom performance and foreign language reading anxiety measured by the three-factor model of the FLRAS. However, this study newly demonstrated a strong association of reading-anxiety subcomponents with learners' reading proficiency. The administration, scoring, and interpretation methods of the FLRAS were reconsidered based on the replicated results.

1. Background

1.1 Foreign Language Reading Anxiety Scale

Psychological affect is one of the individual differences that determine success in second language (L2) learning. Among the various affective variables, research on foreign language anxiety has grown by demonstrating that 'it can interfere with the acquisition, retention, and production of the new language' (MacIntyre & Gardner, 1991, p. 86; see also Horwitz, 2010; De Costa, 2015; Dörnyei & Ryan, 2015). An initial interest of foreign language anxiety research was to operationalize it as a measurable variable, in which Horwitz, Horwitz, and Cope (1986) developed a psychometric inventory titled the Foreign Language Classroom Anxiety Scale (FLCAS). It is composed of 33 items with a 5-point Likert scale, focusing on anxiety of L2 listening and speaking in the classroom. Furthermore, recent studies have examined the constructs of language-skill-specific anxieties to examine their associations with L2 learning, including reading, writing, listening, and speaking (e.g. Cheng, Horwitz, & Schallert, 1999; Pae, 2013; Cheng, 2017).

Foreign language reading anxiety is the worry and negative affectivity that is aroused during reading comprehension in a foreign language (Saito, Horwitz, & Garza, 1999). Since print exposure is an essential source of language input in L2 learning (Grabe, 2009), a better understanding of learners' anxiety in reading should support L2 learning as well as foreign language reading pedagogy. For example, previous studies revealed the negative effects of foreign language reading anxiety on reading comprehension (Sellers, 2000; Rai, Loschky, Harris, Peck, & Cook, 2011; Zhao, Guo, & Dynia, 2013), reading attitude (Yamashita, 2007; Bahmani & Farvardin, 2017), reading strategy use (Lien, 2016), and

incidental L2 vocabulary learning (Zhao, Guo, Biales, & Olszewski, 2016). To quantify this construct as a researchable variable, these studies used the FLRAS developed by Saito et al. (1999). Presently, their study has been cited more than 800 times in Google Scholar; thus, L2 researchers agree that the FLRAS is a standard inventory of foreign language reading anxiety.

When using a scale to examine learners' affects, it is important to assure its psychometric validity, namely, the psychometric soundness of an inventory concerning to what extent it measures what it is designed to measure (Dörnyei, 2003, 2007). The FLRAS has been used in a unidimensional way; thus, a single construct of reading anxiety is measured by its total score. However, Matsuda and Gobel (2004) showed that the constructs measured by the FLRAS were composed of three different causes of foreign language reading anxiety. Their results suggest that it is not appropriate to add up individual response scores to obtain a total score, but rather to divide them to obtain average scores for the three subcategories. Therefore, the present study aimed to resolve the discrepancy in the administration, scoring, and interpretation methods of the FLRAS among previous studies by determining whether Matsuda and Gobel's results would be reproduced. Given the vast number of studies that have referred to the FLRAS when researching foreign language reading anxiety, a replication study will contribute to supporting the validity of the previous studies as well as the original.

1.2 The primary studies

1.2.1 Saito et al. (1999)

Research on foreign language reading anxiety started by exploring the phenomenon of foreign language reading anxiety as distinct from general foreign language anxiety. Saito et al. (1999) assumed that the construct of the FLRAS was different from that of the FLCAS. They hypothesized that foreign language reading anxiety would be caused by two subcomponents: unfamiliar scripts and writing systems, and unfamiliar cultural material. To measure these constructs, the FLRAS included 20 items with a 5-point Likert scale which asked learners to manifest the following psychological affects (Saito et al., 1999, pp. 205–207):

- Anxiety over various aspects of reading (e.g. 'I get upset when I'm not sure whether I understand what I am reading in English').
- Perceptions of reading difficulties in a target language (e.g. 'When reading English, I often understand the words but still can't quite understand what the author is saying').
- Perceptions of the relative difficulties of reading as compared to those of other language skills (e.g. 'The hardest part of learning English is learning to read').

There were two research questions (RQs) related to the psychometric validation of the FLRAS. One was to examine whether foreign language reading anxiety existed as a phenomenon distinguishable from general foreign language anxiety. The other was to test construct validity by exploring whether learners' difficulty in foreign language reading could predict their levels of foreign language reading anxiety. English-speaking students who learned French ($n = 192$), Japanese ($n = 114$), and Russian ($n = 77$) at university answered each statement in both the FLRAS and FLCAS.

As for the first question, a Pearson product-moment correlation analysis was used to test discriminant validity between the FLRAS and FLCAS. Since the two measures shared only 41% of the variance (i.e. $r = .64$, $R^2 = .41$), they concluded that 'there is at least preliminary support for the existence of [foreign language] reading anxiety as a phenomenon related to, but distinct from, general [foreign language] anxiety' (Saito et al., 1999, p. 211). Regarding the second question, the participants were classified into three groups based on their perceptions of the difficulty in their target language. The result of an analysis of variance indicated that learners who perceived reading in their target language as relatively difficult had significantly higher levels of reading anxiety. They also showed that students with higher final grades in a reading course had significantly lower reading anxiety. Thus, foreign language reading anxiety had a negative effect on learners' classroom performance as measured by their

final grades in a reading course. These results supported a part of the psychometric validity of the FLRAS.

Some further studies successfully reproduced this negative relationship between reading anxiety and reading comprehension (e.g. Sellers, 2000; Rai et al., 2011; Zhao et al., 2013), while others could not (e.g. Mills, Pajares, & Herron, 2006; Wu, 2011). The most critical problem of Saito et al. (1999) was that they did not consider the factorial dimensionality of the FLRAS into its scoring and interpretation methods (Hsiao, 2002; Matsuda & Gobel, 2004; Zoghi & Alivandivafa, 2014; see also Park, 2014; Horwitz, 2016). They theoretically supposed that the FLRAS consisted of three sub-categories by distinguishing the questionnaire statements. Thus, it should be inappropriate to add up individual response scores to obtain a total score for the entire inventory; it is necessary to confirm the factor structure of the FLRAS to validly measure learners' foreign language reading anxiety. Development and validation studies on psychometrics have emphasized the importance of considering a substructure of instruments when examining the relationships between learner characteristics and their language skills (Dörnyei, 2003, 2007; Horwitz, 2016; see also Pae, 2013; Baghaei, Hohensinn, & Kubinger, 2014; Park, 2014; Zoghi & Alivandivafa, 2014; Cheng, 2017).

1.2.2 Hsiao (2002)

Hsiao (2002) examined the factorial validity of the Chinese version of the FLRAS because foreign language reading anxiety tends to vary with target L2 languages. This study is not a replication study of Saito et al. (1999); however, it employed a rigorous methodology to investigate the FLRAS's factor structure. In the survey, Hsiao presented the FLRAS to 1,252 Chinese learners of English as a foreign language (EFL) and conducted a factor analysis (FA) with a maximum likelihood method. As the result of an exploratory factor analysis (EFA) with a subsample, one-factor and two-factor models were constructed. Three-factor or four-factor models were not established because at least one latent factor was loaded by only one item. A confirmatory factor analysis (CFA) with the other subsample was conducted to examine which of the one-factor and two-factor models were statistically valid. After eliminating nine problematic items, the two models were compared from the perspective of goodness of fit statistics (see details in Tabachnick & Fidell, 2014). The result showed that the one-factor model produced the most acceptable fit statistics. Table 1 shows Hsiao's one-factor model of the FLRAS.

This study's limitation is that it examined only the factorial validity of the FLRAS. To ensure that the scores obtained from the inventory predict learners' levels of foreign language reading anxiety, further studies need to collect as much evidence as possible to support its construct validity. Specifically, it is necessary to provide evidence of negative associations between foreign language reading anxiety and proficiency in a similar manner as in Saito et al. (1999). The other concern is that the FLRAS was treated as a 6-point Likert scale instead of a 5-point Likert scale. Given that a procedural difference is one of the variables affecting further replications (Porte, 2012), replicability of the factorial structure of the FLRAS has not yet been examined.

1.2.3 Matsuda and Gobel (2004)

Matsuda and Gobel (2004) used the FLRAS to examine reading anxiety factors that could accurately predict classroom performance with 252 Japanese EFL learners. This study submitted the observed data into a principle component analysis (PCA) to extract the factor structure of the FLRAS in an exploratory way. The results, summarized in Table 1, showed that the FLRAS consisted of three sub-components: (a) familiarity with vocabulary and grammar, (b) reading confidence and enjoyment, and (c) familiarity with different cultures, ideas, and writing systems, labeled as language distance. After deleting three items due to problematic factor loadings, the three factors accounted for 41% of the total variance. The most influential factor was familiarity with vocabulary and grammar (21%), followed by reading confidence and enjoyment (12%), and language distance (8%). Thus, this study suggested that the factor structure of the FLRAS was a three-factor model, not a unidimensional or two-factor model.

Table 1 Comparisons of the factor loadings obtained in Hsiao (2002), Matsuda and Gobel (2004), and the present study (PCA and EFA)

Item number and statements	Hsiao (2002)	Matsuda and Gobel (2004)			PCA (The present study)			EFA (The present study)		
	F1	F1	F2	F3	F1	F2	F3	F1	F2	F3
07. When reading English, I get nervous and confused when I don't understand every word.	.74	−.56			.79 [○]	−.14	−.12	.65	−.09	−.16
03. When I'm reading English, I get so confused I can't remember what I'm reading.	.72		−.48		.57	.23	.26	.65	.10	.16
01. I get upset when I'm not sure whether I understand what I am reading in English.	.69	−.59			.71 [○]	−.06	−.23	.62	.05	−.12
06. I get upset whenever I encounter unknown grammar when reading English.	.76	−.74			.55 [○]	−.04	−.01	.60	−.09	−.03
04. I feel intimidated whenever I see a whole page of English in front of me.	.78			−.48	.53	−.31	.05	.52	−.26	.00
05. I am nervous when I am reading a passage in English when I am not familiar with the topic.	.76	−.38			.51 [○]	.19	.30	.47	.03	.20
08. It bothers me to encounter words I can't pronounce while reading English.	.74	−.66			.64 [○]	−.07	.32	.45	.04	.27
02. When reading English, I often understand the words but still can't quite understand what the author is saying.	.62				−.13	.03	.06	.31	−.12	.05
09. I usually end up translating word by word when I'm reading English.	.78			−.63	.09	−.07	.04	.28	−.19	.12
18. I am satisfied with the level of reading ability in English that I have achieved so far. *			.63		.14	.02	.26	−.22	.12	.11
14. Once you get used to it, reading English is not so difficult. *			.63		−.16	.68 [○]	−.15	−.08	.62	−.07
13. I feel confident when I am reading in English. *			.85		−.19	.48 [○]	−.08	−.16	.62	.12
12. I enjoy reading English. *			.74		−.05	.76 [○]	−.15	.09	.60	−.04
17. I don't mind reading to myself, but I feel very uncomfortable when I have to read English aloud.				−.47	−.05	−.03	−.04	−.12	−.36	.24
16. I would be happy just to learn to speak English rather than having to learn to read as well. *					.05	−.26	.47	.29	.33	−.04

15. The hardest part of learning English is learning to read.				.12	.51	−.09	−.15	−.24	−.11
19. English culture and ideas seem very foreign to me.		−.63	−.02	−.01	.74 [○]	.00	.01	.73	
11. I am worried about all the new symbols you have to learn in order to read English.	.69	−.63	−.08	−.21	.65 [○]	−.10	−.28	.41	
10. By the time you get past the funny letters and symbols in English, it's hard to remember what you're reading about.	.76	−.51	.21	−.17	.57 [○]	.19	−.30	.32	
20. You have to know so much about English history and culture in order to read English.	−.42		.05	−.04	.09	−.03	−.01	.15	
Proportion of variance explained by each factor (%)	83.0	20.1	11.8	8.1	14.4	9.4	9.9	14.2	5.9
Cumulative variance explained by each factor (%)	83.0	20.1	31.9	40.0	14.4	23.8	33.7	14.2	29.5

Note. F1 = familiarity with vocabulary and grammar (reading difficulty), F2 = reading confidence and enjoyment, F3 = language distance. The statements with an asterisk were inverted items. The eigenvalues of more than .30 (bolded) were regarded as the corresponding statements contributed to each factor. The eigenvalues with a circle showed that they were composed of the same factor as Matsuda and Gobel (2004).

The other analysis was conducted to predict classroom performance by foreign language reading anxiety measured by the FLRAS. A multiple regression analysis was conducted where the dependent variables were two types of final class grades evaluated in content-based ($n = 89$) and four-skill classes ($n = 89$). The independent variables were two FLCAS subcomponents, three FLRAS subcomponents, gender, and general EFL proficiency measured by the Test of English for International Communication (TOEIC®). The results were inconsistent with Saito et al. (1999); only one component of the FLCAS (self-confidence in speaking) determined classroom performance. None of the three subcomponents of the FLRAS were associated with Japanese EFL learners' classroom performance.

There are two concerns regarding Matsuda and Gobel (2004) from the viewpoint of psychometric validity. First, their interpretation of the PCA results was not robust because some statements cannot obviously be categorized into an intended factor. For example, Table 1 shows that although Items 4 and 9 belong to Factor 3 (i.e. language distance), these statements should manifest learners' anxiety in language processing in foreign language reading. This kind of discrepancy might be caused by the PCA because it does not require any rationale underlying which questionnaire statements should be associated with which factors (Tabachnick & Fidell, 2014). Second, they did not examine whether the multicomponent of foreign language reading anxiety was related to L2 learners' reading proficiency. They only showed that it was not related to classroom performance. Thus, a negative effect of foreign language reading anxiety on L2 reading ability has not been clarified.

1.3 Approaches to replication and psychometric validation

A first concern in the present replication relates to the different assumptions and results on the factor structure of the FLRAS. Saito et al. (1999) hypothesized two kinds of causes of foreign language reading anxiety and three kinds of questionnaire statements to be provided for learners. However, previous studies have used a total or average score of 20 items in the FLRAS because Saito et al. stated that the theoretical range of a score is from 20 to 100. While Hsiao (2002) supported this one-factor model by using EFA with item reductions and CFA, Matsuda and Gobel (2004) agreed to the three-factor model as a result of the PCA.

The significant differences in research methodologies between Hsiao (2002) and Matsuda and Gobel (2004) were the sampled populations (Chinese vs Japanese EFL learners) and statistical approach (FA vs PCA). As for the sample populations, Horwitz (2016, p. 74) agreed that 'a different factor structure might be found in a different population of learners.' Although Rodríguez and Abreu (2003) indicated that foreign language anxiety was stable for a group of highly motivated learners, Kim (2009) and Park (2014) showed that the components of foreign language anxiety varied in sampled populations with respect to instructional, cultural, and proficiency differences. These findings require a replication study to use the same population as far as possible to assess the replicability of the original study.

Methodological differences influence research replicability (e.g. Porte, 2012; Marsden, Morgan-Short, Thompson, & Abugaber, 2018). Regarding the statistical approaches, the PCA with orthogonal rotation applied by Matsuda and Gobel (2004) is based on the assumption that there are no correlations between latent factor variables. The use of FA with oblique rotation indicates that Hsiao (2002) and others (e.g. Pae, 2013; Baghaei et al., 2014; Zoghi & Alivandivafa, 2014; Cheng, 2017) have considered the significant correlations between the latent factors of foreign language reading anxiety. Given that the practical purposes of EFA and CFA are theory development and testing, respectively (Tabachnick & Fidell, 2014), the present study will retain Matsuda and Gobel's methodology as far as possible (i.e. PCA) and additionally use other statistical approaches (i.e. EFA and CFA).

A second goal of this replication is to reexamine the relationship between reading anxiety measured by the FLRAS and classroom performance and reading proficiency. Theoretically, the FLRAS scores have to be associated with reading proficiency to support its psychometric validity. While this rationale has been supported by previous studies (e.g. Sellers, 2000; Rai et al., 2011; Zhao et al., 2013) as well as

Saito et al. (1999), it has also been rejected (Mills et al., 2006; Wu, 2011). If the three-factor model is replicated, one aspect to be raised through replication is to examine what causes of foreign language reading anxiety determine learners' classroom performance and reading proficiency.

Thus, the present replication addressed Matsuda and Gobel's (2004, p. 24) RQ: 'What variables and anxiety factors can be accurate predictors of overall classroom performance?' Since they did not posit an RQ to explicitly state that they examined the number of latent variables the FLRAS included, we investigated their original RQ with respect to (a) the factorial structure of the FLRAS (RQ1) and (b) the relationship between foreign language reading anxiety and classroom performance (RQ2). Unlike the original study, we further examined the negative associations between foreign language reading anxiety and proficiency to confirm the psychometric validity of the FLRAS (RQ3). This allows us to support the potential reliability and generalizability of the conclusions provided by the original studies (Porte, 2012).

RQ1. What kinds of factorial structure of the FLRAS can be reproduced with approximately the same population as in Matsuda and Gobel (2004)?

RQ2. What variables and anxiety factors can be accurate predictors of overall classroom performance?

RQ3. What variables and anxiety factors can be accurate predictors of overall reading performance?

2. Method

2.1 Research design

Table 2 displays the research design of Matsuda and Gobel (2004) and that of the present study. Based on Marsden et al.'s (2018) meta-analysis, this study specified two types of changes for replication as follows: (a) motivated changes and (b) acknowledged but not motivated changes. A motivated change involved conducting EFA and CFA to reproduce Matsuda and Gobel's results.

One of the acknowledged but not motivated changes was the measurements of reading proficiency evaluated by TOEIC® vs TOEIC Bridge® and classroom performance evaluated by end-of-term tests. This study used TOEIC Bridge® for accurate measurements because the participants were not familiar with TOEIC® but were familiar with TOEIC Bridge®. Different course work between the original and present studies caused differences in the specifications of classroom assessments; however, both studies were based on skill-based activities in a classroom including listening, speaking, reading, and writing. While these might affect the reproduced results, the results of this study could be compared with those of the original study because the constructs quantified by those measurements did not differ theoretically from each other. It was also possible to strengthen the potential generalizability of the original study. Because the present study introduced these principled changes from the original study, we labeled it as an approximate (or partial) replication (Porte, 2012).

Marsden et al. (2018) showed that principled changes to participant demographics in L2 replication studies included participants' first language (L1), targeted L2, proficiency, and age. The participants' L1, targeted L2, and age range of the present study were identical to that of participants in Matsuda and Gobel (2004). Although they did administer TOEIC® Listening & Reading to measure their participants' EFL proficiency, they did not report any descriptive statistics. Instead, they labeled participants as high beginner to high intermediate (p. 25).¹ Nevertheless, we did not regard the participants' proficiency levels as substantially different because the level of textbooks used in the original (*Interchange Book 2: A2-level*, Cambridge University Press) and present (*Four Corners Book 2: A2-level*, Cambridge University Press) studies was the same as defined in the Common European Framework for Reference of Languages. Other variables related to the participants' characteristics, such as the sex ratio of the group and their majors, were inevitably different, but were not motivated changes for this replication.

Table 2 Differences between Matsuda and Gobel (2004) and the present study

Variables		Matsuda and Gobel (2004)	The present replication
Participants	<ul style="list-style-type: none">• Number• Age• Sex ratio^b• Major^b• Proficiency	<ul style="list-style-type: none">• 252 Japanese undergraduates• 18 to 21• 177 females, 75 males• English• High beginner to high intermediate	<ul style="list-style-type: none">• 165 Japanese undergraduates• 18 to 21• 23 females, 142 males• Engineering• High beginner to high intermediate
Materials	<ul style="list-style-type: none">• Questionnaire• Reading test^b• Class grade^b	<ul style="list-style-type: none">• FLRAS (Saito et al., 1999)• FLCAS (Horwitz et al., 1986)• TOEIC®• Original test	<ul style="list-style-type: none">• FLRAS (Saito et al., 1999)• TOEIC Bridge®• Original test
Procedure	<ul style="list-style-type: none">• Survey	<ul style="list-style-type: none">• In-class questionnaire	<ul style="list-style-type: none">• In-class questionnaire
Analysis	<ul style="list-style-type: none">• FA^a• Regression	<ul style="list-style-type: none">• PCA (<i>n</i> = 236)• Varimax rotation• Two multiple regressions (<i>n</i> = 89 in four-skill class <i>n</i> = 89 in content-based class)	<ul style="list-style-type: none">• PCA (<i>n</i> = 158)• Varimax rotation• EFA and CFA (<i>n</i> = 158)• Oblimin rotation• Multiple regression (<i>n</i> = 158)

Note.
^aMotivated changes.
^bAcknowledged but not motivated changes.

2.2 Participants

This replication survey’s participants were 165 Japanese EFL learners (female = 23, male = 142, age = 18–21), who were first-year undergraduates majoring in engineering. They had studied English for at least six years in formal education in Japanese secondary schools. Their proficiency ranged from high beginner to high intermediate according to their TOEIC Bridge® scores (*M* = 108.45, *SD* = 16.14, *Min* = 74, *Max* = 156). They were notified of how the data would be used and provided informed consent before answering the questionnaires. They were enrolled in two obligatory 90-minute English courses at university. As mentioned in Table 2, it should be noted that the participants’ sex ratio and major differed between the present study and Matsuda and Gobel (2004).

2.3 Materials

Following Matsuda and Gobel’s (2004) procedure, the questionnaire statements, adopted from Saito et al. (1999: pp. 205–207), were translated into Japanese and discussions were held to check whether the original meanings were retained. The FLRAS (see Table 1) consisted of 20 self-report items as follows: (1) *strongly disagree*, (2) *disagree*, (3) *neither agree nor disagree*, (4) *agree*, and (5) *strongly agree*. The sequence of the statements was rearranged using a random-number method.

2.4 Procedure

The present study followed the data collection procedures used in Matsuda and Gobel (2004). The questionnaire survey was conducted during regular English classes. The participants were asked not to answer the questions based on the specific class in which the questionnaire was administered. Although we did not set any time limit, the administration time was approximately 10 minutes.

The English reading proficiency test (TOEIC Bridge®) was implemented at the beginning of the academic year, and measured participants' English listening and reading skills ranging from 10 to 90, respectively (Maximum possible score = 180, Cronbach's $\alpha = .86$). Classroom performance was evaluated using the average scores of four end-of-quarter tests (Cronbach's $\alpha = .74$). The tests consisted of integrated reading-to-write task performance (50%), independent listening skills (15%), independent reading skills (15%), and spoken interaction (20%).

3. Results and discussion

3.1 RQ1: The factor structure of the FLRAS

Table 3 reports descriptive statistics of the FLRAS responses. The final sample size was 158 after removing missing values by list-wise deletion ($n = 7$). The means, 95% confidence intervals, and standard deviations showed no ceiling or floor effects. The full range of possible values were observed except for Item 18. Normality among the observed variables was assessed in terms of skewness and kurtosis, and univariate and multivariate normality were violated ($Z_s > 1.96$). This result required the use of a robust maximum likelihood method in EFA and CFA (Schmitt, 2011). The Kaiser-Meyer-Olkin measure of sampling adequacy² was satisfied (.75). Bartlett's test of sphericity rejected a null hypothesis of no correlations between the observed variables, $\chi^2(190) = 665.64$, $p < .001$; no multicollinearities were found (r range = $-.36$ to $.47$). These preliminary analyses assured the appropriateness of using PCA, EFA, and CFA. All analyses were performed on R with the packages psych (Revelle, 2018) and lavaan (Rosseel, 2012).

To replicate the factor structure of the FLRAS, we conducted PCA with a varimax rotation method following Matsuda and Gobel (2004). Table 1 shows the factor-loading result of a replicated three-factor model. The model explained 14.4%, 9.4%, and 9.9% of the variances of the observed scores for Factors 1, 2, and 3, respectively. However, the individual statements corresponding to each factor were not identical to Matsuda and Gobel's results. Only 11 out of 17 statements were accurately categorized into the same factors suggested in the original study (agreement ratio = 64.7%).

An alternative statistical approach was used to reexamine the factor structure of the FLRAS. The present study applied an EFA with the robust maximum likelihood and Oblimin rotation methods to the sample ($n = 158$). A parallel analysis³ indicated that the three-factor solution was best. The EFA result showed that each statement was almost appropriately classified into the intended factors (see Table 1); Factors 1, 2, and 3 stood for familiarity with vocabulary and grammar, reading confidence and enjoyment, and language distance, respectively, as suggested in Matsuda and Gobel (2004). According to the grouped items, Factor 1 did not represent only familiarity with vocabulary and grammar, but should be labeled as learners' perception of reading difficulty because it included both lower-level (e.g. Item 7: 'I get nervous and confused when I don't understand every word') and higher-level (e.g. Item 5: 'I am nervous when I am reading a passage in English when I am not familiar with the topic') cognitive processing (Grabe, 2009). This is consistent with Saito et al. (1999), who defined it as anxiety over various aspects of reading. Given that some statements were not properly classified into the intended factors in Matsuda and Gobel, it is necessary to use FA to duplicate the factor structure of the FLRAS.

CFAs were conducted to reexamine the goodness of fit statistics for Hsiao's (2002) one-factor model, Matsuda and Gobel's (2004) three-factor model, and the present three-factor model. First, comparing the goodness of fit statistics in Table 4, the one-factor model showed the second-best fit to the observed data when deleting Items 12 to 20 in the same way as in Hsiao's study (see Figure 1). The item statements comprising the one-factor model reflected learners' perceptions of reading difficulty, because eight out of 11 items are shared with Factor 1 which was elicited by the present study.⁴ Thus, when applying the one-factor model in researching foreign language reading anxiety by the FLRAS, it cannot measure reading confidence and enjoyment (Factor 2) and familiarity with different cultures, ideas, and writing systems (Factor 3).

Table 3 Descriptive statistics of each item response

Items	<i>M</i>	95% CI	<i>SD</i>	<i>Min</i>	<i>Max</i>	Skewness	Kurtosis	<i>SE</i>
Item 1	3.74	[3.60, 3.88]	0.92	1	5	−0.71	0.41	0.07
Item 2	3.56	[3.38, 3.73]	1.13	1	5	−0.61	−0.57	0.09
Item 3	3.08	[2.91, 3.24]	1.03	1	5	−0.22	−0.92	0.08
Item 4	3.90	[3.73, 4.07]	1.10	1	5	−0.95	0.22	0.09
Item 5	3.18	[3.01, 3.35]	1.09	1	5	−0.13	−0.88	0.09
Item 6	3.56	[3.39, 3.73]	1.08	1	5	−0.75	−0.28	0.09
Item 7	3.70	[3.53, 3.88]	1.10	1	5	−0.71	−0.28	0.09
Item 8	2.77	[2.60, 2.94]	1.09	1	5	0.14	−1.03	0.09
Item 9	3.16	[2.98, 3.34]	1.16	1	5	−0.09	−0.92	0.09
Item 10	2.82	[2.66, 2.99]	1.05	1	5	0.16	−0.72	0.08
Item 11	2.99	[2.82, 3.17]	1.10	1	5	0.33	−0.96	0.09
Item 12	2.90	[2.73, 3.07]	1.08	1	5	−0.07	−0.77	0.09
Item 13	1.95	[1.80, 2.10]	0.94	1	5	1.12	1.15	0.08
Item 14	3.06	[2.89, 3.23]	1.10	1	5	−0.13	−1.02	0.09
Item 15	2.82	[2.66, 2.97]	1.01	1	5	0.08	−0.77	0.08
Item 16	3.48	[3.31, 3.65]	1.08	1	5	−0.34	−0.45	0.09
Item 17	2.97	[2.78, 3.17]	1.27	1	5	−0.01	−1.21	0.10
Item 18	1.59	[1.48, 1.70]	0.71	1	4	1.22	1.57	0.06
Item 19	2.34	[2.19, 2.48]	0.91	1	5	0.37	−0.13	0.07
Item 20	3.01	[2.84, 3.17]	1.05	1	5	−0.01	−1.07	0.08

Note. *n* = 158.

Table 4 Goodness of fit statistics and model selection indices among three models

Model	Goodness of fit statistics					Model selection	
	CFI	TLI	RMSEA	90% CI	SRMR	AIC	BIC
Hsiao (2002)	.93	.91	.05	[.01, .07]	.07	4,945	5,013
Matsuda and Gobel (2004)	.77	.73	.07	[.05, .08]	.09	7,529	7,525
The present study	.95	.94	.05	[.03, .06]	.06	3,448	3,529

Matsuda and Gobel’s (2004) model showed the least best fit to the data. Their model was established using PCA on the assumption that any factors would not be correlated with each other. However, Figure 2 shows that reading difficulty (Factor 1), reading confidence and enjoyment (Factor 2), and language distance (Factor 3) were strongly associated. Although this appears to be empirically (Yamashita, 2007) and theoretically (Grabe, 2009) consistent with past studies, the most critical problem is positive associations between Factor 2 and Factors 1 and 3. Whereas Factor 2 reflects self-confidence in reading, Factors 1 and 3 reflect anxiety in reading and different cultures; therefore, the two different factors have to be correlated negatively. This may be one of the reasons for the weak association between foreign language reading anxiety and classroom performance in their study.

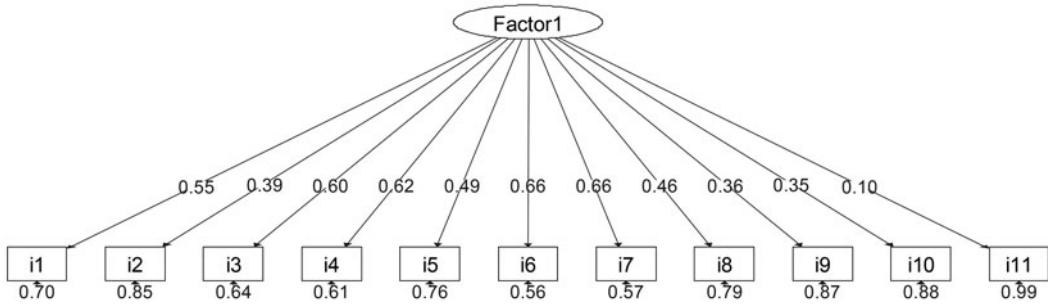


Figure 1. The replicated one-factor model of the FLRAS suggested by Hsiao (2002). All the standardized path coefficients except Item 11 were significant at the level of $p < .05$.

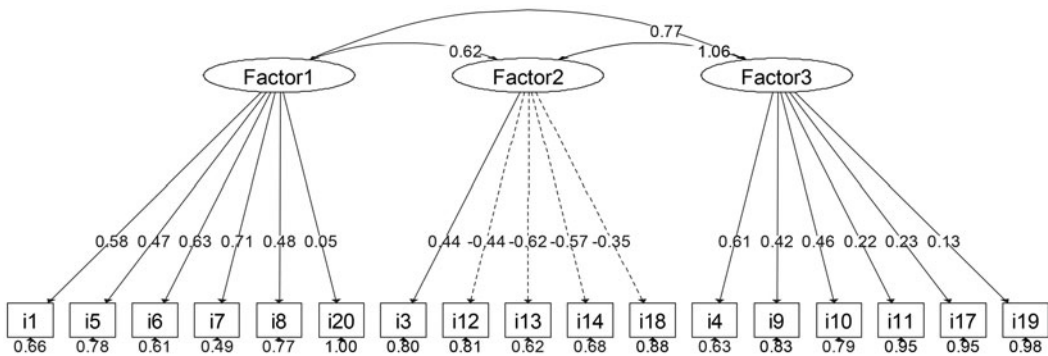


Figure 2. The replicated three-factor model of the FLRAS suggested in Matsuda and Gobel (2004). All the standardized path coefficients except Items 19 and 20 were significant at the level of $p < .05$.

As discussed above, the other perspective for Matsuda and Gobel's (2004) problematic factor model is the differences in the item statements comprising each factor. The present study reached the same conclusion as Matsuda and Gobel in that the three-factor solution was best for the FLRAS; however, our model produced the best fit to the data. The model selection indices (Akaike's Information Criterion [AIC] and Bayesian Information Criterion [BIC]) showed that the revised three-factor model was the most statistically plausible. As Figure 3 shows, the associations between each factor are acceptable because of a positive relationship between Factors 1 and 3 and a negative relationship between Factor 2 and Factors 1 and 3. This suggests that the FLRAS can measure three different constructs of foreign language reading anxiety.

3.2 RQ2: Predictors of classroom performance

Table 5 shows means with 95% confidence intervals, standard deviations, and intercorrelations for dependent and independent variables. A multiple regression analysis examined what causes of foreign language reading anxiety affected classroom performance. Following Matsuda and Gobel (2004), the independent variables included participants' overall proficiency estimated by TOEIC Bridge® and gender, and three factors of reading anxiety.

Table 6 showed that Factor 3 predicted class grades, indicating that anxiety in language distance suppressed classroom performance ($\beta = -.22$, $p = .003$). This was inconsistent with Matsuda and Gobel (2004), who showed that no factors were related to Japanese EFL learners' classroom

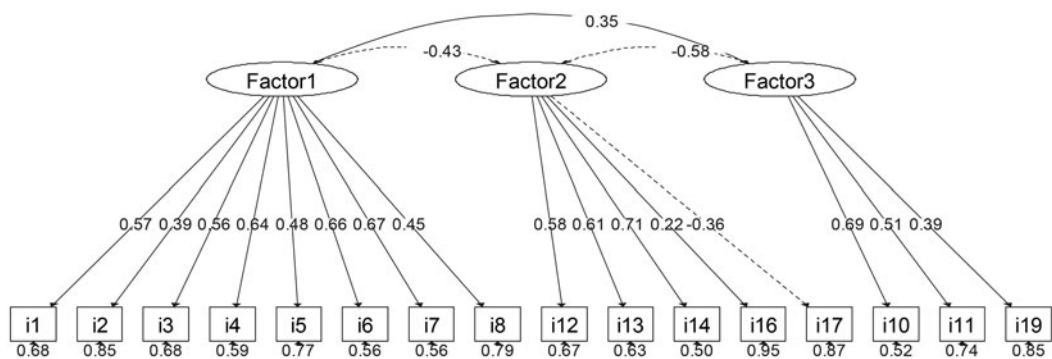


Figure 3. The three-factor model of the FLRAS suggested by the present study. All the standardized path coefficients were significant at the level of $p < .05$.

Table 5 Means, standard deviations, and intercorrelations for the measurements

Variables	M	95% CI	SD	1.	2.	3.	4.	5.
Class grades	81.85	[80.59, 83.12]	8.02	-.10	.13	-.28*	.48*	.10
Predictors								
1. Factor 1	3.42	[3.31, 3.53]	0.69	—	-.31*	.04	-.26*	-.04
2. Factor 2	2.85	[2.74, 2.96]	0.71		—	-.17*	.37*	-.03
3. Factor 3	2.66	[2.54, 2.79]	0.82			—	-.17*	-.10
4. Proficiency	108.45	[105.90, 110.99]	16.14				—	.12
5. Gender	n.a.	n.a.	n.a.					—

Note. Factor 1 = reading difficulty, Factor 2 = reading confidence and enjoyment, Factor 3 = language distance.
* $p < .05$.

Table 6 Results of multiple regression analysis with the FLRAS for classroom performance

Predictors	<i>B</i>	95% CI	<i>SE B</i>	β	<i>t</i>	<i>p</i>
Factor 1	0.11	[-1.07, 1.30]	0.60	.01	0.19	.850
Factor 2	-0.45	[-1.70, 0.79]	0.63	-.06	-0.72	.472
Factor 3	-1.75	[-2.89, -0.62]	0.57	-.22	-3.05	.003
Proficiency	0.22	[0.15, 0.30]	0.04	.45	5.82	<.001
Gender	0.47	[-2.76, 3.69]	1.63	.02	0.28	.776
(Intercept)	57.43	[49.16, 65.71]	4.19		13.72	<.001

Note. Factor 1 = reading difficulty, Factor 2 = reading confidence and enjoyment, Factor 3 = language distance. $F(5, 152) = 10.92$, $p < .001$. $R^2 = .27$.

performance. The measurements for classroom performance differed between the original and present studies. While the differences may produce inconsistent outcomes, the statements consisting of the three anxiety factors differed between the two studies. For example, Matsuda and Gobel's Factor 3 included Items 4 ('I feel intimidated whenever I see a whole page of English in front of me'), 9 ('I usually end up translating word by word when I'm reading English'), and 17 ('I don't mind reading

to myself, but I feel very uncomfortable when I have to read English aloud'), which were categorized into either Factors 1 or 2 in the present study. Since it was unlikely for these statements to reflect learners' anxiety in learning unfamiliar scripts, writing systems, and cultural material (i.e. Factor 3), revised Factor 3 (Items 10, 11, and 19) was a significant predictor of classroom performance. The present result also supports Saito et al. (1999) and Zhao et al. (2013) who showed the negative effect of language distance on classroom reading performance.

The present study partly reproduced Matsuda and Gobel's (2004) result by finding no associations between learners' classroom performance and their reading difficulty ($\beta = .01$, $p = .850$) and reading confidence and enjoyment ($\beta = -.06$, $p = .472$). Saito et al. (1999) and previous studies (e.g. Sellers, 2000; Rai et al., 2011; Zhao et al., 2013) showed that students who received higher grades manifested significantly lower levels of overall reading anxiety than those with lower grades. According to Matsuda and Gobel, these inconsistent outcomes can be attributed to differences in the participants' course work. For example, Saito et al. provided students with instructions specific to reading skills such as word reading, strategy use, and text comprehension.⁵ However, Matsuda and Gobel and the present study implemented listening, speaking, and writing in addition to reading. Matsuda and Gobel evaluated their students' classroom performance by writing/reading assignments and grammar-based tests (50%), class participation (40%), and attendance (10%) in the four-skills class. They also used writing assignments and grammar-based tests (20%) and oral presentation (80%) in the content-based class. Similarly, the percentage of independent reading measures in the present study was relatively small (15%). Thus, whether the FLRAS is able to predict learners' classroom performance depends on the degree of engagement in reading in a class. Assuming that Saito et al.'s class grades mainly reflected their students' reading performance achieved through the course work, these discrepant results should be resolved if the FLRAS scores were associated with learners' reading proficiency.

3.3 RQ3: Predictors of reading proficiency

Descriptive statistics of the TOEIC Bridge® reading score were as follows: $M = 52.78$ (95% CI [51.19, 54.37]), $SD = 10.10$. This score was weakly correlated with Factor 1 ($r = -.27$, $p = .001$), Factor 2 ($r = .37$, $p < .001$), and Factor 3 ($r = -.22$, $p = .006$). Table 7 shows the results of a multiple regression analysis; the dependent variable was the reading proficiency score and the independent variables were three causes of foreign language reading anxiety and gender.

Unlike the regression model for classroom performance, all three factors were significant predictors of reading proficiency. The three factors explained 19% of the variance in reading proficiency. Reading comprehension was inhibited by a higher difficulty in reading processing ($\beta = -.17$, $p = .032$), less confidence in reading ($\beta = .30$, $p < .001$), and higher anxiety in language distance ($\beta = -.15$, $p = .040$). These results demonstrate that the psychological affects measured by the FLRAS are related to foreign language reading proficiency, but not to general classroom performance. Matsuda and Gobel (2004) showed that the FLCAS better predicts classroom performance than the FLRAS. They did not examine the relationships between the three causes of reading anxiety and reading proficiency; however, the present findings provide evidence for the psychometric validity of the FLRAS.

Although the percentage of variance explained by Hsiao's (2002) one-factor model ($R^2 = .14$) was lower than the present three-factor model ($R^2 = .19$), the reading difficulty was also an independent predictor of reading performance ($B = -6.54$, 95% CI [-9.14, -3.94], $SE B = 1.32$, $\beta = -.37$, $t = 4.96$, $p < .001$). Previous studies produced contradictory findings by adopting a one-factor model by summing up almost all the statements of the FLRAS to examine the associations between reading anxiety and proficiency. Since the one-factor model (i.e. Items 1 to 11) could predict participants' reading performance, we support those negative relationships (Sellers, 2000; Rai et al., 2011; Zhao et al., 2013) and reject the null results (Mills et al., 2006; Wu, 2011) as long as the FLRAS is appropriately employed.

Table 7 Results of multiple regression analysis with the FLRAS for reading proficiency

Predictors	<i>B</i>	95% CI	<i>SE B</i>	β	<i>t</i>	<i>p</i>
Factor 1	−1.68	[−3.21, −0.15]	0.77	−.17	−2.17	.032
Factor 2	3.00	[1.44, 4.56]	0.79	.30	3.80	<.001
Factor 3	−1.55	[−3.03, −0.07]	0.75	−.15	−2.07	.040
Gender	1.80	[−2.40, 6.01]	2.13	.06	0.85	.398
(Intercept)	52.50	[50.94, 54.07]	0.79		66.34	<.001

Note. Factor 1 = reading difficulty, Factor 2 = reading confidence and enjoyment, Factor 3 = language distance. $F(4, 153) = 9.16$, $p < .001$. $R^2 = .19$.

4. Conclusions

4.1 Summary and limitations

This approximate replication sought to assure the psychometric validity of the FLRAS with respect to its factorial structure and relationships between reading anxiety and proficiency. Matsuda and Gobel's (2004) results could not be replicated well when using the same statistical approaches. Instead, our study successfully reproduced the three-factor model using EFA and CFA (RQ1) and the weak relationship between reading anxiety and classroom performance (RQ2). Although the original study did not explore the relationships between the three factors of reading anxiety and reading proficiency, the present study supported the theoretical constructs of the FLRAS by demonstrating the inhibition effects of reading anxieties on reading comprehension (RQ3). Thus, when the use of the FLRAS is appropriate, this inventory can safely be used in a classroom and in foreign language reading anxiety research.

Nevertheless, this study's results should be viewed in the context of research limitations because they affect the administration, scoring, and interpretation of the FLRAS. First, the questionnaire responses were collected from Japanese EFL university students with the aim of replicating the results of Matsuda and Gobel (2004). However, the restricted range in terms of L1 background, targeted L2, proficiency, and age may be considered for further psychometric validation to generalize our conclusions and those of Matsuda and Gobel (Marsden et al., 2018). Psychological affects including foreign language anxiety are sometimes context-dependent (e.g. Kim, 2009; Zhao et al., 2013; De Costa, 2015; Dörnyei & Ryan, 2015) and some researchers have developed their own instruments targeting a specific population (e.g. Baghaei et al., 2014; Zoghi & Alivandivafa, 2014). If the results of this study cannot be reproduced with different populations, it is necessary to conduct similar replication studies using context-specific instruments with corresponding populations.

In addition, the present study could not exactly replicate Matsuda and Gobel (2004) with respect to the measurements of classroom performance and reading proficiency. Because it is difficult to use the exact same measurements in different intact classes, research synthesis is necessary focusing on the extent to which foreign language reading anxiety is associated with classroom performance. For example, Teimouri, Goetze, and Plonsky's (2019) meta-analysis has indicated that foreign language anxiety is moderately associated with course grade ($k = 33$, $r = -.34$, 95% CI [−.39, −.29]). To this end, they have emphasized the importance of reporting the grade specifications such as measurement source, descriptive statistics, and reliability indices.

4.2 Administration

The FLRAS can be administered individually (Rai et al., 2011) and to groups of L2 learners (Saito et al., 1999). Although there is no time limit, the estimated administration time ranges from five to ten minutes. After explaining the purpose of the questionnaire, teachers need to ask students to read each statement and choose the closest scale to their thoughts and perceptions. Teachers should

check that students understand all aspects of the inventory. It is important to tell students that their responses do not refer to the specific class where the questionnaire is administered (Matsuda & Gobel, 2004) and do not contribute to their class grade (Dörnyei, 2003).

4.3 Scoring

The valid way to calculate the anxiety scores is to average the scores provided by the three subscales, such as reading difficulty (Items 1, 2, 3, 4, 5, 6, 7, and 8), reading confidence and enjoyment (Items 12, 13, 14, 16, and 17), and language distance (Items 10, 11, and 19). The reverse code scale items (Items 12, 13, 14, and 16) must be corrected before summing and averaging the scores so that a high value manifests the same type of response on the other items. While the present study employed the factor scores in the psychometric validation, the correlations between the factor scores and the averaged scores were extremely high (r range = .98–1.00). This indicates that both scores can be used alternatively. Using a total score for the entire inventory is not recommended unless adopting Hsiao's (2002) one-factor model (Items 1 to 11) for the measurements.

4.4 Interpretation

The average scores for each subscale in the FLRAS indicate the causes of foreign language reading anxiety in L2 learners. Currently, there are no cut-off points to interpret the magnitude of their anxiety, unlike Oxford's (1990) Strategy Inventory for Language Learning.⁶ The average scores should be used to determine whether learners have relatively higher anxiety than is caused by any other source. It is necessary to note relatively high scores on any of the subscales or specific statements because specific anxieties may interfere with learning to read in a foreign language (Saito et al., 1999). However, dealing with foreign language anxiety will depend on the situational contexts of a classroom, such as individual differences in L2 learners (e.g. Kim, 2009; Horwitz, 2010; De Costa, 2015; Lien, 2016), their reading proficiency (e.g. Saito et al., 1999; Yamashita, 2007; Rai et al., 2011; Zhao et al., 2013), text difficulty (Bahmani & Farvardin, 2017), and other potentially related factors (see Dörnyei & Ryan, 2015). Great care should be taken when using the present results in score interpretation for different populations from Japanese EFL university students.

Endnotes

¹ We thank Professor Sae Matsuda for acknowledging that Matsuda and Gobel's (2004) data has been lost because their study was conducted prior to the development of efficient data storage mediums.

² Values of .60 and above are required for good FA (Tabachnick & Fidell, 2014).

³ Simulation research has indicated that the parallel analysis is the best empirical method for selecting the appropriate number of factors (Schmitt, 2011). In the analysis, a series of data sets generated using the sample size and number of variables of the original data set is used to examine whether real non-random factors exist.

⁴ Two of the 11 statements (Items 10 and 11) reflect learners' familiarity with different writing systems.

⁵ Although Saito et al. (1999) reported the use of cumulative scores on daily quizzes, lesson quizzes, oral exams, final exams, and homework assignments to determine students' classroom performance, they did not specify the degree to which the scores independently reflected reading performance.

⁶ Oxford's (1990) Strategy Inventory for Language Learning categorizes the magnitudes of strategy use as high (mean of 3.5 or higher), medium (mean of 2.5 to 3.4), and low (mean of 2.4 or lower) based on the questionnaire responses.

Acknowledgements. This research was supported by Japan Society for the Promotion of Science Grants-in-Aid for Young Scientists (B) No. 17K13512 and 18K12443. We thank Professor Sae Matsuda, the author of the original study, for her cooperation. We would also like to thank Dr Graeme Porte and the *Language Teaching* reviewers for their insightful comments on earlier versions of this article.

References

- Baghaei, P., Hohensinn, C., & Kubinger, K. D. (2014). Persian adaptation of foreign language reading anxiety scale: A psychometric analysis. *Psychological Reports*, 114(2), 315–325.

- Bahmani, R., & Farvardin, M. T. (2017). Effects of different text difficulty levels on EFL learners' foreign language reading anxiety and reading comprehension. *Reading in a Foreign Language*, 29(2), 185–202.
- Cheng, Y.-S. (2017). Development and preliminary validation of four brief measures of L2 language-skill-specific anxiety. *System*, 68, 15–25.
- Cheng, Y.-S., Horwitz, E. K., & Schallert, D. L. (1999). Language anxiety: Differentiating writing and speaking components. *Language Learning*, 49(3), 417–446.
- De Costa, P. I. (2015). Reenvisioning language anxiety in the globalized classroom through a social imaginary lens. *Language Learning*, 65(3), 504–532.
- Dörnyei, Z. (2003). *Questionnaires in second language research: Construction, administration, and processing*. New York: Routledge.
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. New York: Oxford University Press.
- Dörnyei, Z., & Ryan, S. (2015). *The psychology of the language learner revisited*. New York: Routledge.
- Grabe, W. (2009). *Reading in a second language: Moving from theory to practice*. New York: Cambridge University Press.
- Horwitz, E. K. (2010). Foreign and second language anxiety. *Language Teaching*, 43(2), 154–167.
- Horwitz, E. K. (2016). Factor structure of the Foreign Language Classroom Anxiety Scale: Comment on Park (2014). *Psychological Reports*, 119(1), 71–76.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *Modern Language Journal*, 70(2), 125–132.
- Hsiao, T.-Y. (2002). Unidimensionality of the Chinese version of the Foreign Language Reading Anxiety Scale. *Perceptual and Motor Skills*, 95(3), 927–933.
- Kim, S.-Y. (2009). Questioning the stability of foreign language classroom anxiety and motivation across different classroom contexts. *Foreign Language Annals*, 42(1), 138–157.
- Lien, H.-Y. (2016). Effects of EFL individual learner variables on foreign language reading anxiety and metacognitive reading strategy use. *Psychological Reports*, 119(1), 124–135.
- MacIntyre, P. D., & Gardner, R. C. (1991). Methods and results in the study of anxiety and language learning: A review of the literature. *Language Learning*, 41(1), 85–117.
- Marsden, E., Morgan-Short, K., Thompson, S., & Abugaber, D. (2018). Replication in second language research: Narrative and systematic reviews and recommendations for the field. *Language Learning*, 68(2), 321–391.
- Matsuda, S., & Gobel, P. (2004). Anxiety and predictors of performance in the foreign language classroom. *System*, 32(1), 21–36.
- Mills, N., Pajares, F., & Herron, C. (2006). A reevaluation of the role of anxiety: Self efficacy, anxiety, and their relation to reading and listening proficiency. *Foreign Language Annals*, 39(2), 276–295.
- Oxford, R. L. (1990). *Language learning strategies: What every teacher should know*. Boston: Heinle & Heinle.
- Pae, T.-I. (2013). Skill-based L2 anxieties revisited: Their intra-relations and inter-relations with general foreign language anxiety. *Applied Linguistics*, 34(2), 232–252.
- Park, G.-P. (2014). Factor analysis of the Foreign Language Classroom Anxiety Scale in Korean learners of English as a foreign language. *Psychological Reports*, 115(1), 261–275.
- Porte, G. (2012). *Replication research in applied linguistics*. New York: Cambridge University Press.
- Rai, M. K., Loschky, L. C., Harris, R. J., Peck, N. R., & Cook, L. G. (2011). Effects of stress and working memory capacity on foreign language readers' inferential processing during comprehension. *Language Learning*, 61(1), 187–218.
- Revelle, W. R. (2018). psych: Procedures for Personality and Psychological Research [R package]. Evanston. Retrieved from <https://personality-project.org/r/psych-manual.pdf>
- Rodríguez, M., & Abreu, O. (2003). The stability of general foreign language classroom anxiety across English and French. *Modern Language Journal*, 87(3), 365–374.
- Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48(2), 1–36.
- Saito, Y., Horwitz, E. K., & Garza, T. J. (1999). Foreign language reading anxiety. *Modern Language Journal*, 83(2), 202–218.
- Schmitt, T. A. (2011). Current methodological considerations in exploratory and confirmatory factor analysis. *Journal of Psychoeducational Assessment*, 29(4), 304–321.
- Sellers, V. D. (2000). Anxiety and reading comprehension in Spanish as a foreign language. *Foreign Language Annals*, 33(5), 512–520.
- Tabachnick, B. G., & Fidell, L. S. (2014). *Using multivariate statistics (6th edn)*. Edinburgh: Pearson Education Limited.
- Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, 41(2), 363–387.
- Wu, H.-J. (2011). Anxiety and reading comprehension performance in English as a foreign language. *Asian EFL Journal*, 13(2), 273–307.
- Yamashita, J. (2007). The relationship of reading attitudes between L1 and L2: An investigation of adult EFL learners in Japan. *TESOL Quarterly*, 41(1), 81–105.

- Zhao, A., Guo, Y., & Dynia, J. (2013). Foreign language reading anxiety: Chinese as a foreign language in the United States. *Modern Language Journal*, 97(3), 764–778.
- Zhao, A., Guo, Y., Biales, C., & Olszewski, A. (2016). Exploring learner factors in second language (L2) incidental vocabulary acquisition through reading. *Reading in a Foreign Language*, 28(2), 224–245.
- Zoghi, M., & Alivandivafa, M. (2014). EFL reading anxiety inventory (EFLRAI): Factorial validity and reliability. *Journal of Psychoeducational Assessment*, 32(4), 318–329.

Akira Hamada (Ph.D., University of Tsukuba) is a lecturer in the Department of English, Faculty of Languages and Cultures at Meikai University, Japan. His research interests include second language reading, instructed second language acquisition, and educational assessment.

Shuichi Takaki (Ph.D., University of Tsukuba) is an associate professor in the Faculty of Human Development and Culture at Fukushima University, Japan. His research interests include second language reading, language assessment, and research methodology.

Cite this article: Hamada, A., & Takaki, S. (2019). Approximate replication of Matsuda and Gobel (2004) for psychometric validation of the Foreign Language Reading Anxiety Scale. *Language Teaching* 1–17. <https://doi.org/10.1017/S0261444819000296>