

Reading anxiety scales: Do they measure the same construct?

Hitoshi Mikami
Chubu University
Japan

Abstract

This article discusses the interchangeability of three self-report measures for reading anxiety. Despite their differences in target constructs, the three scales have been used for similar lines of research. After computing shared variance between the target scales and examining the behavior of anxiety indexes in relation to the amount of graded reading and reading comprehension performance, the author reaches the conclusion that each of the three scales should be treated as a unique research tool (i.e., the scales are non-interchangeable with one another). The overall results also suggest the utility of each reading anxiety scale in educational research and practice.

Keywords: reading anxiety, self-report questionnaire, scale selection, graded reading, reading comprehension performance

Self-report questionnaires have been a prevalent tool for measuring anxiety in second and foreign language (L2 and FL) research (Cheng, 2017; Dewaele & MacIntyre, 2016; Pae, 2013). The degree of reading anxiety measured through such self-reporting is found to have a relationship with, for example, reading performance (Mikami, Leung, & Yoshikawa, 2018; Sellers, 2000; Zhao, Guo, & Dynia, 2013), reading proficiency (Yamashita, 2007), and educational treatments (Bahmani & Farvardin, 2017; Yamashita, 2013). There is, however, a fundamental problem in the current use of reading anxiety scales: while several different scales were used in similar lines of research, it remains untested whether existing reading anxiety scales are basically interchangeable in empirical research or we should treat each scale as a unique research tool. If the latter is the case, it is necessary to start considering how scales should be selected in future studies. To obtain a clear picture on this issue, this study examined the behavior of anxiety attributes measured through three pre-existing scales in two separate pieces of research. Recommendations for the future application of the anxiety scales in educational research and practice are provided based on these observations.

A Brief Review of Reading Anxiety Scales

Otto, Calkins, and Hearon (2010) defined anxiety as “an emotional state characterized by a sense of apprehension, worry, and lack of control of one’s own affective response” (p.131). Accordingly, reading anxiety refers to a specific type of anxiety aroused by the elements tied to

the reading of additional language, such as unknown words in texts, unfamiliar topics, imperfect text comprehension, and a lack of self-confidence in reading (Brantmeier, 2005; Cheng, 2017; Saito, Garza, & Horwitz, 1999; Yamashita, 2007; Young, 1999). Similar to many other psychological constructs, self-report questionnaires have been a common tool for measuring reading anxiety. In terms of the number of citations and the use of these questionnaires in recent studies, there are three popular anxiety questionnaires in the domain of reading. They are, the scale introduced in Saito et al. (1999), that used in Yamashita (2007, 2013), and the original and modified scale of Brantmeier (2005).¹ To clarify each scale's target construct and their contribution to reading research, in what follows the author provides a quick review of the three scales.

FL Reading Anxiety Scale (FLRAS)

Saito et al. (1999) was the first study to introduce a scale specifically designed to measure anxiety toward reading in an additional language. As we can expect from a large number of question items comprising the FLRAS ($k = 20$), Saito and colleagues capture reading anxiety as a multidimensional construct (see the Appendix for all 20 question items). To be specific, Saito et al.'s (1999) framework posits that the following factors determine the level of general reading anxiety: (a) nervousness and worry associated with imperfect text comprehension (e.g., encounter with unknown linguistic features), (b) negative or pessimistic beliefs about reading (e.g., holding the idea that reading is the hardest part of language learning), (c) lack of enjoyment or self-confidence in reading, (d) unwillingness to demonstrate one's linguistic skill (e.g., nervousness about reading target language aloud in classes), and (e) reliance on first language (L1) in FL reading.

The primary significance of the FLRAS is in its sensitivity to measure domain-specific anxiety. Saito and associates' original study first demonstrated that the FLRAS measures something different from general FL anxiety: they observed 59% of non-shared variance between the FLRAS and the index for general FL anxiety (for similar findings, see, Joo & Damron, 2015; Matsuda & Gobel, 2001). The discriminant validity of the FLRAS was further backed by a later study, as Pae (2013) confirmed the independence of the FLRAS from other skill-based language anxiety scales.

The results of Saito et al. (1999) further supported the validity argument of the FLRAS: in their observation, (a) native English speakers studying Japanese language (i.e., a typical non-Western language) had higher reading anxiety compared to their counterparts studying Western languages, and (b) the level of reading anxiety increased together with perceived difficulty in FL reading. The overall result reported in Saito et al. (1999) not only made the FLRAS a popular tool in FL and L2 reading research (Bahmani & Farvardin, 2017; Joo & Damron, 2015; Matsuda & Gobel, 2004; Zhao et al., 2013), but demonstrated the benefit of using general reading anxiety (i.e., the multidimensional construct) in empirical studies.

Brantmeier's Scale and its Adaptation

Taking into account pre-existing anxiety scales (Horwitz, Horwitz, & Cope, 1986; Saito et al., 1999; Young, 1999), Brantmeier (2005) developed a 10-item anxiety scale (i.e., Brantmeier's

scale). Of the 10 items, seven target reading anxiety, encompassing apprehension about reading activities (e.g., reading for homework) and that about oral and written tasks during or after reading (e.g., reading L2 aloud in classes and writing compositions about what was read). The results of Brantmeier (2005) were informative in that (a) the behavior of reading anxiety was unique from that concerning productive skills (measured through two of the remaining items), and (b) even advanced L2 readers in her study increased their anxiety level in situations where a post-reading task (either oral or written) came together with a reading activity.

As a limitation, however, Brantmeier (2005) leaves open the question of whether an increase in reading anxiety affects the performance of a reading task. This bottleneck problem was overcome when Mikami et al. (2018) tested their hypothesis. Namely, that the influence of reading anxiety on reading performance becomes systematically negative in settings where an intense cognitive effort is required for task completion (see Zhao, Guo, & Dynia, 2013 for a similar discussion). A modified version of the Brantmeier's scale (the MBS) was used for their testing. The MBS ($k = 4$) focuses primarily on apprehension about (a) reading itself, (b) the reading of lengthy texts, and (c) performance in post-reading tasks (see the Appendix for the question items). The test results upheld the hypothesis on the negative influence of reading anxiety on reading performance ($R^2 = .25$). The possibility that emerges from the results of Brantmeier (2005) and Mikami et al. (2018) is that an amalgam of worry about reading and that about task performance plays a central role in the elevation of reading anxiety in task situations.

Yamashita's Anxiety Scale (YAS)

As part of her research on reading attitude, Yamashita (2007) identified four items that form a single factor called Anxiety. What the YAS is mainly concerned with is worries about general reading comprehension (see the Appendix for the question items). Using the YAS, Yamashita (2007) demonstrated that 21% of variance in Anxiety was explained by reading proficiency measured by TOEIC[®]. Also, in her subsequent study, Yamashita (2013) demonstrated that the implementation of an extensive reading program (i.e., a program designed to minimize failure in reading comprehension with the help of graded reading materials) is a practical way to decrease the YAS score ($r = .34$) (for a similar result, see Banno & Kuroe, 2016) and increase the feeling of comfort toward L2 reading ($r = .54$).

When Yamashita's (2007, 2013) results are integrated with the findings that reading anxiety explains variance in reading performance (Mikami et al., 2018; Sellers, 2000, Zhao et al., 2013), a bidirectional relationship between reading anxiety and reading skill development is conceptualized: less proficient readers are at a higher risk of failure in reading comprehension; unsuccessful reading attempts strengthen negative feelings toward reading, encompassing anxiety (Bahmani & Farvardin, 2017) and discomfort; and the negative feelings manifest themselves as disruptive anxiety in reading task events. This rationale first allows reading anxiety and proficiency (measured by linguistic tasks) to explain each other's variance, and, more importantly, it is open to the possibility that reading anxiety is treatable with the provision of optimal learning experience.

Functional Difference among Reading Anxiety Scales: A Research Gap

As previously discussed, the theoretical constructs underlying each one of the three reading anxiety scales are at least in part different. It is easy, however, to find cases where different reading anxiety scales are used for similar lines of research. First, both the FLRAS and YAS were used to discuss the influence of reading programs on reading anxiety (Bahmani & Farvardin, 2017; Banno & Kuroe, 2016; Yamashita, 2013), and the same goes for the clarification of the relationship between reading anxiety and learning outcomes (Matsuda & Gobel, 2004; Yamashita, 2008). Also, both the FLRAS and Brantmeier's scale (including the MBS) were employed for investigating the anxiety–performance interface (Brantmeier, 2005; Joo & Damron, 2015; Mikami et al., 2018; Zhao et al., 2013). If the three scales are interchangeable with one another, the scale selection should not create any practical problems. Meanwhile, if the opposite is the case, it is necessary to start considering the appropriate usage of the three reading anxiety scales in empirical research.

The Current Research

To obtain a clear picture on the aforementioned issues, this study examined the following research questions (RQs). RQ1 concerns to what degree the three anxiety scales tap into the same latent dimension. RQ2 and 3 deal with the issue of how the score of each anxiety scale behaves in research designs where more than one reading anxiety scale has been employed in the past.

1. To what degree do the three anxiety scales (the FLRAS, YAS, and Brantmeier's scale or MBS) share variance with each other?
2. How do the three reading anxiety scales (listed in RQ1) correlate with the performance of a cognitively demanding reading comprehension task?
3. How do the three reading anxiety scales (listed in RQ1) correlate with the amount of graded reading?

Method

Research Design

The data collection was conducted two separate times. The first took place from January to February 2018, recruiting 66 English majors (English L2 learners) at a Japanese university, and the second in October 2018, targeting another 57 English majors. All participants spoke Japanese as their L1 and belonged to the same department. There was a possibility, therefore, that the data of all recruits are applicable to the computation of shared variance between the anxiety indexes (RQ1). The remaining RQs were answered using the data of either one group or the other, because different criteria were applied to recruitment for each. The first 66 students were recruited for RQ2 (i.e., the anxiety–performance relationship). The condition that they were within the same L2 proficiency range allowed the execution of a cognitively demanding task for all participants (Mikami et al., 2018; Zhao et al., 2013). Regarding RQ3 (i.e., the relationship between anxiety and graded reading), the 57 recruits were chosen because they have been taking

a graded reading program in 2018, while the aforementioned 66 students finished the same program at least a year before the recruitment. To prevent the subgrouping leading to insufficient sample sizes for statistical testing, the minimal numbers of recruits were determined with reference to the p -values and effect sizes of previous studies (see the next section for details). All recruits gave permission to use their data for the current research project. The procedure undertaken in this study also met the ethical requirements of the institution involved.

General Procedure

RQ2: The anxiety–performance relationship. All students of the target department were required to take TOEIC® on a regular basis and the author was entitled to access their latest test score; thus, the sampling started by dividing all potential recruits into several proficiency groups. The grouping was conducted by relating the TOEIC® reading scores to the levels of the Common European Framework of Reference (CEFR) (Educational Testing Service, 2013). The number of minimal recruits was set at 40 referring to the results of Mikami et al. (2018) ($N = 35$). In the end, the 66 students in the CEFR A2 (i.e., Waystage) level were recruited. All 66 recruits first responded to a questionnaire survey. The survey documented the recruits' background information (e.g., name, gender, age, L1) and reading anxiety scores. Two to three weeks after the survey, 58 out of 66 initial recruits worked with a reading material for the B1 level (i.e., a level higher than their current level—A2), and received 2,000 Japanese Yen for their cooperation.²

RQ3: Anxiety and graded reading. The 57 students were recruited for two reasons. The first is the fact that they had been taking a graded reading program at the time of the data collection. The second was the sample size close to those of previous studies, in which the statistical link between anxiety and graded reading was observed (Bahmani & Farvardin, 2017; Yamashita, 2013). The 57 students completed the questionnaire survey, and submitted their activity data in the reading program.

Indexes

Reading proficiency. The participants' latest TOEIC® reading section score (possible score range: 5–495) represents their English reading proficiency (for the same indexing, see Yamashita, 2007). The TOEIC® reading section consists of 100 questions and estimates reading proficiency based on the five subskill scores (Cid, Wei, Kim, & Hauck, 2017).

Reading anxiety. The FLRAS ($k = 20$), MBS ($k = 4$), and YAS ($k = 4$) were employed for the measurement of reading anxiety (see the Appendix for the question items). The MBS was selected over the original scale due to (a) its empirical relationship with the reading performance of Japanese L1 English learners (Mikami et al., 2018), and (b) the necessity for testing the reproducibility of the MBS–performance link across settings. Also, provided they are supported by some sort of validity, short version questionnaires contribute to efficient data collection. Regarding the YAS, as with Yamashita (2013), the group of four items was treated as the measurement of the single latent variable (i.e., Anxiety), because the sample size of this study was insufficient for the use of confirmatory factor analysis (Phakiti, 2018). All 28 question items were translated into Japanese and the target language was changed into English. The

intelligibility of translation was checked by 10 university students who did not appear in the data collection procedure. The questions were answered on a five-point Likert scale (ranging from “strongly disagree” to “strongly agree”).

Reading comprehension. This study employed a written recall protocol for the documentation of reading comprehension. The participants at the CEFR A1 level first read a 451-word factual story for the B1 level (i.e., a level higher than A2) and then wrote down whatever they could remember from the text. The recall was done in the participants’ L1 so that English writing skill did not cause task-irrelevant variance in the recall score (Brantmeier, Strube, & Yu, 2014). To accustom them to the recall procedure, the participants completed a preparation task (i.e., a recall task with a simple two-line text) before task engagement. The reading material used here was an excerpt from Cambridge English: Preliminary entitled *Wildlife Filmmakers*.³

Supplemental data. After completing the task, the participants answered three questions about the text they had just read. The first question asked if the participants had ever read *Wildlife Filmmakers* somewhere before (a yes-no question). The second question recorded to what degree the participants were familiar with the topic discussed in the text (a four-point Likert scale: 1 = *Not familiar at all* and 4 = *Very familiar*). The last question documented the perceived difficulty of the text (a six-point Likert scale: 1: *Very easy* and 6: *Very hard*).

The amount of graded reading. This index represents the total number of words read at week 20 of the graded reading program. The program was a four-quarter design (i.e., 7 weeks \times 4), and the reading amount at week 20 was used in this study because (a) the anxiety survey for the 57 students took place on the last day of that week, and (b) it was not possible to conduct an anxiety survey after week 20 because of the students’ class schedule. The program was equipped with all five principles of an effective program as reported by Jeon and Day (2015). First, the reading activity was individual and silent. Second, over 2,200 books available in the program gave the students freedom in the selection of books. Third, to borrow a book, the participants logged in to an online system *M-reader*, checked grade-level lists, and chose a book suited to their reading level. Levels started from zero and increased once a participant had finished reading three books of a given level. The next principle, “learners read as much as possible” (Jeon & Day, 2015, p.302) was achieved by (a) informing the students the benefits of graded reading (Day et al., 2015), and (b) setting a minimal goal of 25,000 words per quarter. Each time a participant failed to reach this goal, ten percentage points were deducted from the final grade of their English class. The *M-reader* system monitors the participants’ achievement using a quiz on a chosen book (in most cases $k = 10$). The number of words in a book were counted as one’s achievement when the number of correct answers about its content exceeded a predetermined threshold. There were three chances to pass the quiz part per book, and no feedback on the quiz was given upon failure.⁴ Fifth, teachers gave necessary advice to those who did not perform well during the program execution. The reading activity was done outside of English classes, and all participants were taking the same English language modules outside the reading program.

Data Analysis

Final Sample Size

No outlier was identified regarding the anxiety data (the Tukey's method). Meanwhile, the data of the two participants in the recall task were excluded from further analysis because their recall consisted of information completely irrelevant to the reading material (i.e., $N = 56$ for RQ2, the anxiety-comprehension relationship). Regarding RQ3 (the relationship between anxiety and graded reading), the data of all 57 participants were retained as no questionable behavior, such as sudden spikes in reading amount just before a deadline, was detected in the *M-reader* record. Only the data of one group ($N = 56$) was used for the calculation of shared variance between the anxiety indexes (RQ1) due to the difference in the two groups' responses to the YAS. Specifically, the Cronbach's α of the 57 participants ($\alpha = .55$) was not only well outside the acceptable range ($\alpha \leq .70$), but much lower than that observed in Yamashita (2013) ($\alpha = .68$ and $.72$). To deal with this issue, the author decided to (a) use the data of the 56 participants (the YAS's $\alpha = .69$) for RQ1, and (b) discuss the reason for the different responses based on the results of the data analysis.

Reading Anxiety Indexes

Both the MBS and YAS consist of four items. The target of the MBS is worry about reading and comprehension performance in a post-reading task, and that of the YAS is apprehension about general reading comprehension. Regarding the FLRAS, the author prepared four types of indexes in order to be able to observe the scale's behavior in as much detail as possible. The first was an index consisting of all 20 items, which represents overall anxiety toward reading (hereafter this specific index will be referred to as the FLRAS). The rest of the indexes were adopted from Matsuda and Gobel (2004), which classified the FLRAS into three partially overlapping factors (i.e., the FLRA subscales) sampling 252 English majors at a Japanese university. The sample sizes of this study ($N = 56$ and 57) were insufficient to conduct a confirmatory analysis for the factor structure (Phakiti, 2018). The same grouping of items as in Matsuda and Gobel (2004) was thus used, except where item trimming led to greater homogeneity in the items at the current setting (Grabowski & Oh, 2018). Two cultural items (item 7 and 15 in the Appendix) were abandoned for this reason. As a result, the number of items was six ($k = 6$) for all three FLRA subscales (see the Appendix for the grouping). As with Matsuda and Gobel (2004), the Grammar & Vocabulary index measures anxiety toward the linguistic features of English and that of Confidence/Enjoyment measures the lack of self-confidence and enjoyment in English reading (see p.28 of the original study). Language Distance was redefined as the perceived linguistic distance between L1 and English, due to the removal of the cultural item (i.e., item 15 in the Appendix).

RQ1: Anxiety Scales and Overlap in Their Target Constructs

To confirm to what degree the three anxiety scales tap into the same latent dimension, this study computed shared variance (r^2) between the MBS, YAS, and FLRAS and its subscales using the partial correlation coefficients ($N = 56$). The control variable was the TOEIC[®] score because reading proficiency is expected to affect one's anxiety estimates (Yamashita, 2007).

RQ2: Anxiety and Reading Comprehension Performance

The post-task survey clarified that no participant had previously read *Wildlife Filmmakers*. To calculate the recall score (i.e., the degree of text comprehension), the author first divided the reading material into 75 pausal units. That is, the points where a native speaker takes a natural pause (i.e., inserts brackets) when they read L1 texts out loud at a normal speech rate (for the validity of this analytical method, see Brantmeier et al., 2014). The bracketing was done by a native English speaker (a bachelor in Linguistics) and the appropriateness of pausing was double-checked by another native English speaker. The correct recall of one pausal unit was given one point except for pauses that had no pragmatic meaning when used alone (i.e., “as”). This made the possible recall score range between 0.00 and 74.00. The scoring was done by the author and done on two occasions eight weeks apart, between the assessments. The intraclass correlation coefficient of the scoring was satisfactory ($= .96$).

This study performed Spearman’s partial rank correlation analysis to confirm the relationship between the anxiety indexes and performance in a demanding reading comprehension task (i.e., the recall score) ($N = 56$). The Spearman’s test was selected because the skewness of some indexes was more than twice the standard error of skewness (see Table 1 in the section for results). The author controlled for the influence of topic familiarity, perceived task difficulty, and reading proficiency in the analysis as these elements also caused performance difference in the task. The internal consistency of the anxiety indexes was acceptable for the FLRAS ($\alpha = .79$), Grammar & Vocabulary ($\alpha = .72$), Language Distance ($\alpha = .70$), the MBS ($\alpha = .85$), and marginally so regarding the YAS ($\alpha = .69$). Confidence/Enjoyment ($\alpha = .56$) was removed from the analysis because item trimming did not improve the homogeneity of the items.

RQ3: Anxiety and the Amount of Graded Reading

In prior studies, successful reading attempts were observed to provide impetus for reducing reading anxiety (Yamashita, 2013), while the reading of difficult materials did the opposite (Bahmani & Farvardin, 2017). On this point, the 57 participants had read English books suitable to their reading level and the amount of reading in the *M-reader* system did not increase unless readers demonstrated their comprehension of the key points of selected books. The participants with a greater amount of reading under this condition experienced a greater amount of success, in that they made more attempts to understand L2 texts and succeeded in doing so. The relationship hypothesized from this situation is a negative one: namely, the more books the participants read, the less anxious they will be about L2 reading.

This study performed Spearman's rank correlation analysis to confirm the relationship between the graded reading and the anxiety indexes ($N = 57$) (see the standard error of skewness in Table 2). The internal consistency was acceptable for the FLRAS ($\alpha = .73$), Grammar & Vocabulary ($\alpha = .74$), Language Distance ($\alpha = .70$), and the MBS ($\alpha = .81$). Because item trimming did not improve the homogeneity of the items, the YAS ($\alpha = .55$) and Confidence/Enjoyment ($\alpha = .56$) were removed from the analyses.

Statistical Computations

R version 3.1.1 (The R Core Team, 2016) was employed for statistical computations. The standardized index scores and two-tailed tests were used for testing. Alpha was set at $p \leq .050$ using the false discovery rate control procedure (Benjamini & Hochberg, 2000). The Cohen's criteria was applied for the interpretation of effect sizes (Cohen, 1992).

Results

Descriptive Statistics

Table 1 shows the descriptive statistics for RQ1 and 2 ($N = 56$), and those for RQ3 are cited in Table 2 ($N = 57$). The mean anxiety scores in Tables 1 and 2 were clustered around three in the 5-point range, indicating that the participants were moderately anxious about English reading as a group. Next, on average, about 5% ($M = 3.43$, $SD = 1.91$, $Mdn = 3.00$) of the total number of pausal units (i.e., 74 units) was recalled in the reading task.

Table 1. *Descriptive statistics on the eight indexes (raw scores)*

Index	<i>M</i>	95% CI	<i>SD</i>	<i>Mdn</i>	<i>Skew</i>
FLRAS	3.15	[3.01, 3.29]	0.52	3.15	-0.10
Grammar & Vocabulary	3.49	[3.30, 3.69]	0.73	3.50	-0.64
Language Distance	3.21	[3.00, 3.41]	0.76	3.17	0.03
MBS	3.22	[2.95, 3.50]	1.04	3.13	-0.22
YAS	3.49	[3.28, 3.70]	0.79	3.50	-0.34
Topic Familiarity	2.05	[1.69, 2.42]	1.35	1.00	0.64
Perceived Task Difficulty	4.89	[4.68, 5.10]	0.78	5.00	-0.95
Recall Score	3.43	[2.90, 3.95]	1.91	3.00	1.01
Reading Proficiency	187.64	[176.29, 198.98]	41.95	180.00	-0.23

Note. the standard error of skewness is 0.33 for $N = 56$.

Table 2. *Descriptive statistics on the four indexes (raw scores)*

Index	<i>M</i>	95% CI	<i>SD</i>	<i>Mdn</i>	<i>Skew</i>
FLRAS	3.13	[3.01, 3.24]	0.44	3.20	-0.62
Grammar & Vocabulary	3.37	[3.19, 3.56]	0.72	3.50	-1.08
Language Distance	3.08	[2.74, 3.23]	0.93	3.33	-0.45
MBS	2.98	[2.74, 3.23]	0.93	2.75	0.24

Note. the standard error of skewness is 0.32 for $N = 57$.

Figure 1 is the summary of the graded reading index (RQ3). Individual differences in the reading amount were large at week 20 of the program (i.e., a week before the deadline for the minimal goal of 75,000 words). To be specific, the achievement of 16 participants was below the goal by 15,000 words or more ($Min = 49,472$), whereas 13 participants had finished reading 100,000 words or more ($Max = 602,778$).

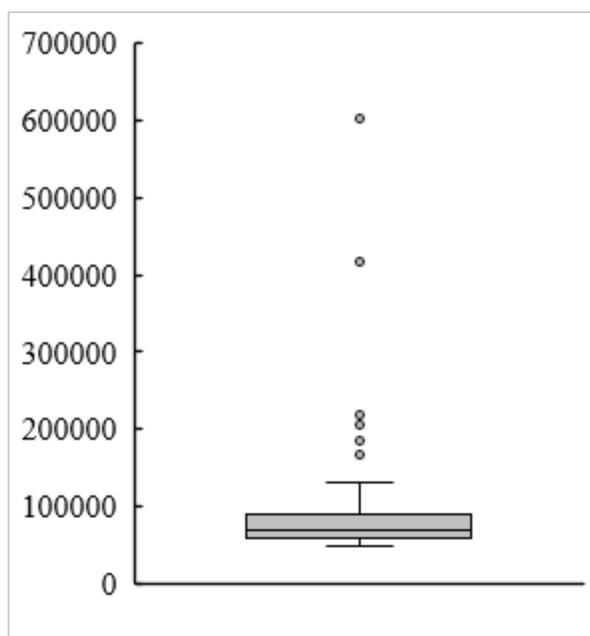


Figure 1. Total number of words read in the graded reading program ($N = 57$), the Tukey's method was used for the illustration of the box plot ($M = 102,272$, $Mdn = 68,662$, $SD = 98,261$, $Min = 49,472$, $Max = 602,778$, $Skew = 3.59$).

RQ1: To What Degree do the Three Anxiety Scales Share Variance with Each Other?

Table 3 shows shared variance between the reading anxiety indexes. When the influence of reading proficiency was controlled, 65% to 67% of variance was shared between the FLRAS and the other two anxiety scales (i.e., the MBS and YAS) while shared variance between the MBS and YAS dropped to 52%. The latter situation (i.e., roughly 50% shared variance) generally applied to the relationship between the FLRA subscales, and MBS and YAS.

Table 3. *Shared variance between the reading anxiety indexes*

Index	Grammar & Vocabulary	Language Distance	MBS	YAS
FLRAS	-	-	.67	.65
Grammar & Vocabulary	-	.60	.50	.63
Language Distance		-	.49	.51
MBS			-	.52

Note. Controlled for the influence of reading proficiency ($df = 53$); all results were significant at the $p > 0.01$ level (corrected by the BH method).

RQ2: How Do the Three Reading Anxiety Scales Correlate with the Performance of a Cognitively Demanding Reading Comprehension Task?

Table 4 summarizes how the reading comprehension index correlated with the anxiety indexes. All five anxiety indexes had a negative relationship with reading performance at the point estimate. Of these correlations, the FLRAS ($r_s = -.31, p = .023$) and the MBS ($r_s = -.28, p = .038$) showed a moderate correlation with the performance in the recall task, even when controlling for the influence of reading proficiency, perceived task difficulty, and topic familiarity. In addition, the correlation between Grammar & Vocabulary and the performance measure approached significance ($r_s = -.27, p = .054$).

Table 4. *Partial correlation (r_s and their 95% CI) between reading anxiety and reading comprehension performance*

Index	FLRAS	Grammar & Vocabulary	Language Distance	MBS	YAS
Recall Score	-.31* [-.05, -.53]	-.27 [-.01, -.50]	-.23 [.04, -.47]	-.28* [-.02, -.51]	-.14 [.13, -.39]

Note. Controlled for the influence of topic familiarity, perceived task difficulty, and reading proficiency ($df = 50$, each); 95% CIs were calculated using the Fisher's r to Z transformation; * = $p \leq .05$ (corrected by the BH method).

RQ3: How Do the Three Reading Anxiety Scales Correlate with the Amount of Graded Reading?

Table 5 shows the results of correlation analysis. Although all correlations in Table 5 were negative at the point estimate, only the FLRAS ($r_s = -.34, p = .005$) and the MBS ($r_s = -.27, p = .050$) had a significant correlation with the amount of words read in the graded reading program. These two correlations were moderate in size, and their 95% CIs did not include zero. These results confirm that the participants who read more in the graded reading program were less anxious in the FLRAS and MBS.

Table 5. Correlation (r_s and their 95% CI) between the anxiety indexes and the total number of words read in the graded reading program

Index	FLRAS	Grammar & Vocabulary	Language Distance	MBS
Amount of Graded Reading	-.34** [-.09, -.55]	-.18 [.08, -.42]	-.19 [.07, -.43]	-.27* [-.01, -.50]

Note. 95% CIs were calculated using the Fisher's r to Z transformation; * = $p \leq .05$, ** = $p \leq .01$ (corrected by the BH method).

Discussion

Reading Anxiety, Recall Performance, and Reading Amount

On average, about 5% of the total number of pausal units was recalled in the reading task. This result is not surprising for two reasons. First, an average of 10 to 30% recall is common in the cases of novice and intermediate L2 readers (Sellers, 2000), and our participants belonged to such a proficiency range (i.e., the CEFR A1 level). Also, recall performance was expected to be below the common results because all 56 participants dealt with a reading task difficult for their L2 level. What is worthy of attention is, therefore, the result indicating that the attribute (or attributes) measured through the FLRAS and MBS was (or were) related to variance in the recall score, even where performance was low.

The minimal requirement of 75,000 words seems to have affected the participants' reading behavior to some extent ($Mdn = 68,662$ words one week before the deadline); at the same time, however, individual differences in this index were quite large ($Min = 49,472$, $Max = 602,778$, $Skew = 3.59$). This, in addition to the system in which text comprehension was key to increasing the reading amount, gives variance in reading amount a practical meaning. Namely, there were differences in the efforts made to read, and in the amount of successful text comprehension. The moderate correlations observed in this study thus uphold the benefit of graded reading on the decrease in reading anxiety (Bahmani & Farvardin, 2017).

Interchangeability of the Three Anxiety Scales

Given that the MBS and YAS consist of a relatively small number of items ($k = 4$, each) and concern specific aspects of reading anxiety (i.e., the MBS: worry about reading and post-reading comprehension tasks; the YAS: worry about general reading comprehension), it is not surprising that 48% of variance was not shared between these indexes. A more important question here is then whether 33 to 35% of non-shared variance between the FLRAS and 4-item scales suggests that the MBS and YAS are only small subsets of the 20-item FLRAS or whether these scales measure something outside the scope of the FLRAS.

This study upholds the latter possibility—each of the three scales measures something different, at least in part. This argument is based on two observations. First, the item-by-item comparison of the scales reveals that some items of the MBS and YAS target unique aspects of reading anxiety. As for the MBS, this exclusively concerns anxiety arousal in a post-reading comprehension task ($k = 1$) and the reading of lengthy texts ($k = 2$). These two attributes thus

appear to be the unique targets of the MBS. Moving our attention to the YAS, one question concerns learners' tendency to worry about a future negative event ("I sometimes feel anxious that I may not understand even if I read."), whereas the FLRAS focuses only on the anxiety levels estimated (or beliefs formed) based on respondents' past experience (e.g., "I get upset whenever I encounter unknown grammar when reading English."). The expectation of failure in reading and negative emotions attached to such a prediction seems to be, therefore, what the YAS uniquely measures.

The second argument stems from the result that the correlation patterns of the FLRAS and MBS are similar in Table 6, despite 33% of non-shared variance between them and the difference in their target attributes. This result demonstrates that the FLRAS and MBS had similar magnitudes of correlation with the target measures but through different pathways. Also, the behavior of the YAS is unique from any other anxiety indexes in Table 6. Although we must be cautious with the interpretation of insignificant results, it is logical to hypothesize that the behavior of the exclusive target of the YAS, the negative expectation for reading comprehension, was different from the other anxiety indexes. What overall observations suggest is that the results of this study should be interpreted focusing on the unique target of each anxiety scale, and this leads us to the view that the FLRAS, MBS, and YAS should be treated as unique research tools instead of interchangeable ones.

Table 6. *Summary of test results*

Index	Reliability	Comprehension Performance	Amount of Graded Reading
FLRAS	.73 – .79	$r_s = -.31^*$	$r = -.34^{**}$
Grammar & Vocabulary	.72 – .74	$r_s = -.27$	$r = -.18$
Language Distance	.70	$r_s = -.23$	$r = -.19$
Confidence/Enjoyment	.56	-	-
MBS	.81 – .85	$r_s = -.28^*$	$r = -.27^*$
YAS	.55 – .69	$r_s = -.14$	-

* = $p \leq .05$, ** = $p \leq .01$ (corrected by the BH method), - = untested due to low Cronbach's alpha.

Future Usage of the Reading Anxiety Scales

This section makes recommendations for the future usage of the reading anxiety indexes based on interpretations of their behaviors in this study. The author will first develop his rationale focusing on the significant test results. Next, the author will introduce the possible adjustments that future studies could take regarding the use of the FLRA subscales and YAS, as it is beyond the ability of a single article to draw a negative conclusion on the validity of measurements.

The practical value of the MBS seems to emerge in situations where research or educational practice involves the reading of lengthy texts and a post-reading comprehension task. First, the reading of graded books (i.e., lengthy texts) and the completion of post-reading quizzes in the *M-reader* system are connected to the unique target of the MBS. Also, while the 56 A2-level

participants were typically capable of understanding short and simple everyday topics in English reading (The Council of Europe, 2001), the skill required in the recall task was the comprehension of a story and its details. This gap shows that the task material would have been perceived as lengthy to the 56 learners. Such a perception, together with the pressure for the upcoming recall task, appears to lead to anxiety arousal during task-taking, and created the link between the MBS and task performance.

Exploratory investigations would particularly benefit from the use of the FLRAS. The findings of prior studies showed that the relative relationship between learners' L1 and their target language (e.g., differences in the writing system and cultural distance) affected the score of some FLRAS items (Saito et al., 1999; Zhou, 2017). Regarding this, the multidimensionality of the FLRAS ($k = 20$) helps identify key attributes behind the increase in reading anxiety in our target learner group, especially when little is known about target learners prior to recruitment for the research. What is more, the results of this study backed the empirical value of the broadly-defined construct of general reading anxiety (i.e., the FLRAS was one of the two anxiety indexes that had clear correlations with both the target measures). The FLRAS is therefore recommended as the first option for data collection when researchers and educators wish to (a) explore the role of general reading anxiety in language learning and use, and (b) determine the key items promoting reading anxiety in a given context.

A natural question arising from the behavior of the FLRAS is whether the application of its subscales is beneficial in capturing the detailed behavior of reading anxiety. Although the results of this study were in favor of the FLRAS being used as a 20-item index, the behavior of Grammar & Vocabulary (that is, worry about the linguistic features of English) gives hints for potential usage of the subscales. This index had, albeit marginally non-significant, a quasi-medium correlation with the performance measure ($r_s = -.27, p = .054$). Saito et al. (1999) added worry about grammar, vocabulary, and unfamiliar topics (i.e., the components of Grammar & Vocabulary) to the FLRAS because language learners often try to have perfect comprehension in reading attempts. Despite such a tendency, however, the 56 participants in the recall task had only a partial understanding of the reading material (i.e., on average, about 5% of the total number of pausal units). As the task takers could hardly be confident of their text comprehension in this situation, the quasi-medium inverse correlation in Tables 4 and 6 can be interpreted as the manifestation of Grammar & Vocabulary. The possibility derived from this rationale is that the FLRA subscales function as a practical tool in situations where the findings of prior studies or the results of a pilot study allow us to form a logical link between the subscale and its target measure for confirmatory investigations.

The behavior of the YAS is worth retesting for two reasons. First, although the YAS's alpha values were marginally adequate in Yamashita (2013) ($\alpha = .68$ and $.72$), $\alpha = .55$ obtained in this study was much lower even than the lower value obtained by Yamashita (2013). As the alpha values of all other five anxiety indexes remained either consistently low or adequate, the structural flaw in the sampling of this study can hardly be the reason for this problem. The data of this study instead suggest that the internal consistency of the YAS may be sensitive to learners' reading proficiency. A basic characteristic of the 56 participants who gave relatively consistent responses to the YAS ($\alpha = .69$) in this study is that they are at the CEFR A2-level. These learners are, again, typically capable of comprehending simple everyday topics in reading

(The Council of Europe, 2001). This result may reflect a situation whereby the learners at the A2 level share the same reasons to be anxious in reading comprehension, whereas individual differences on this issue become larger when learners are less or more proficient than the A2 level. This reasoning suggests that, when the behavior of the YAS is retested, careful attention should be paid to learners' reading proficiency and its effect on the internal consistency of the YAS.

Another possibility is that the graded reading program implemented over a year before the task engagement weakened the YAS's correlation with the comprehension measure ($r_s = -.14$). This argument may seem contradictory to the result that the FLRAS and MBS had medium correlations with the comprehension measure. It is intriguing, however, that the significant correlations observed in this study ($r_s = -.31$ and $-.28$) were weaker than those reported in prior studies, such as those in Zhao et al. (2013) (FLRAS–performance: $r = -.45$ and $-.41$) or that of Mikami et al. (2018) ($r = -.50$). The primary difference between the sample of this study and those of the prior studies is that the latter sampled FL learners with no previous graded reading experience. This difference in sample selection increased the likelihood that the reading program that the 56 participants had undergone mediated anxiety toward English reading (Bahmani & Farvardin, 2017; Yamashita, 2013), and in turn weakened the association between the six anxiety indexes and performance measure. While this rationale strengthens the utility of the FLRAS and MBS in a wide range of learning contexts, it also leaves the possibility that the YAS, the score of which is reportedly sensitive to experience of extensive reading (Yamashita, 2013), may show a greater correlation with the comprehension measure when used in a typical FL context (i.e., a limited amount of language classes and no prior experience of graded reading).

Limitations

Two challenges remain for future studies. The first is the need to retest the behavior of the YAS in a typical FL context. Second, even though 95% CI of the significant correlations in this study did not include zero and this study was moderate in sample size compared to those of previous studies, the provision of more precise parameter estimates on the correlation coefficients is encouraged to make the statistical analysis more robust (Plonsky, 2015). Thus, the possibilities discussed in this study should be tested with a large sample size.

Conclusion

This study upholds the idea that the three reading anxiety scales measure something different to each other at least in part. It thus becomes necessary for future studies to elaborate the logical link between a chosen scale and their target variables. It is my hope that the findings of this study will assist future studies to choose the reading anxiety scale best suited to their research designs.

Acknowledgments

I am grateful to RFL's anonymous reviewers for their informative comments on this manuscript. I would also like to thank Natalie-Anne Hall (The University of Manchester) for English language editing. Note that this work was supported by JSPS KAKENHI Grant Number JP18K12472.

Notes

1. According to Google Scholar, Saito et al. (1999) was cited by 809 studies, Yamashita (2007, 2013) were cited by 188 studies, and Brantmeier (2005) was cited by 132 studies. Note that this literature search was conducted on 3 October 2018.
2. The 58 students also completed extra surveys and dealt with a task different from the one that appears in this study. These extra investigations were conducted for another research project, and thus, these data will be reported separately.
3. One can access the text used in this study from the following website: <https://www.cambridgeenglish.org/exams-and-tests/preliminary/preparation/> (accessed on 15 December 2017). The author was granted permission (via E-mail) from Cambridge Assessment to use their test material for investigational purposes, on 14 December 2017.
4. One can find more information on how the *M-reader* system works at the following website: <https://mreader.org> (accessed on 12 October 2018).

References

- 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*, 185–202.
- Banno, E., & Kuroe, R. (2016). Effects of extensive reading on Japanese language learning. In G. Melanie (Ed.), *Proceedings of the 3rd world congress on extensive reading* (pp. 1–10). Dubai: Extensive Reading Foundation.
- Benjamini, Y., & Hochberg, Y. (2000). On the adaptive control of the false discovery rate in multiple testing with independent statistics. *Journal of Educational and Behavioral Statistics, 25*, 60–83. <https://doi.org/10.3102/10769986025001060>
- Brantmeier, C. (2005). Anxiety about L2 reading or L2 reading tasks? A study with advanced language learners. *The Reading Matrix(2), 5*, 67–85.
- Brantmeier, C., Strube, M., & Yu, X. (2014). Scoring recalls for L2 readers of English in China: Pausal or idea units. *Reading in a Foreign Language, 26*, 114–130.
- Cheng, Y. (2017). Development and preliminary validation of four brief measures of L2 language-skill-specific anxiety. *System, 68*, 15–25. <https://doi.org/https://doi.org/10.1016/j.system.2017.06.009>

- Cid, J., Wei, Y., Kim, S., & Hauck, C. (2017). *Statistical analyses for the updated TOEIC® Listening and Reading test*. New Jersey: Educational Testing Service.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*, 155–159. <https://doi.org/http://dx.doi.org/10.1037/0033-2909.112.1.155>
- Day, R., Bassett, J., Bowler, B., Parminter, S., Bullard, N., Furr, M., ... Robb, T. (2015). *Extensive reading (revised edition)*. Oxford: Oxford University Press.
- Dewaele, J., & MacIntyre, P. (2016). Foreign language enjoyment and foreign language classroom anxiety. The right and left feet of FL learning? In P. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 215–236). Bristol: Multilingual Matters.
- Educational Testing Service. (2013). Mapping the TOEIC® and TOEIC Bridge™ tests on the common European framework of reference for languages. Retrieved November 18, 2017, from http://www.ets.org/s/toeic/pdf/toeic_cef_mapping_flyer.pdf
- Grabowski, K. C., & Oh, S. (2018). Reliability analysis of instruments and data coding. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The palgrave handbook of applied linguistics research methodology* (pp. 541–565). London: Palgrave MacMillan.
- Horwitz, E. K., Horwitz, M. B., & Cope, J. (1986). Foreign language classroom anxiety. *The Modern Language Journal*, *70*, 125–132. <https://doi.org/10.1111/j.1540-4781.1986.tb05256.x>
- Jeon, E.-Y., & Day, R. R. (2015). The effectiveness of core ER principles. *Reading in a Foreign Language*, *27*, 302–307.
- Joo, K. Y., & Damron, J. (2015). Foreign language reading anxiety: Korean as a foreign language in the United States. *Journal of the National Council of Less Commonly Taught Languages*, *17*, 23–55.
- Matsuda, S., & Gobel, P. (2001). Quiet apprehension: reading and classroom anxieties. *JALT Journal*, *23*, 227–247.
- 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>
- Mikami, H., Leung, C. Y., & Yoshikawa, L. (2018). The threshold of anxiety in low-stakes testing for foreign language reading. *Reading in a Foreign Language*, *30*, 92–107.
- Otto, M. W., Calkins, A. W., & Hearon, B. A. (2010). Anxiety. In I. B. Weiner & W. E. Craighead (Eds.), *The Corsini encyclopedia of psychology* (4th ed., pp. 131–133). New Jersey: Wiley.
- Pae, T.-I. (2013). Skill-based L2 anxieties revisited: Their intra-relations and the inter-relations with general foreign language anxiety. *Applied Linguistics*, *34*, 232–252. <https://doi.org/https://doi.org/10.1093/applin/ams041>
- Phakiti, A. (2018). Confirmatory factor analysis and structural equation modeling. In A. Phakiti, P. De Costa, L. Plonsky, & S. Starfield (Eds.), *The palgrave handbook of applied linguistics research methodology* (pp. 459–500). London: Palgrave MacMillan.
- Plonsky, L. (2015). *Advanced quantitative methods in second language research*. New York: Routledge.
- Saito, Y., Garza, T. J., & Horwitz, E. K. (1999). Foreign language reading anxiety. *The Modern Language Journal*, *83*, 202–218. <https://doi.org/10.1111/0026-7902.00016>
- Sellers, V. D. (2000). Anxiety and reading comprehension in Spanish as a foreign language. *Foreign Language Annals*, *33*, 512–520. <https://doi.org/10.1111/j.1944-9720.2000.tb01995.x>

- The Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.
- The R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.r-project.org/>
- Yamashita, J. (2007). The relationship of reading attitudes between L1 and L2: An investigation of adult EFL learners in Japan. *TESOL Quarterly*, 41, 81–105.
<https://doi.org/10.2307/40264332>
- Yamashita, J. (2008). Extensive reading and development of different aspects of L2 proficiency. *System*, 36, 661–672. <https://doi.org/10.1016/j.system.2008.04.003>
- Yamashita, J. (2013). Effects of extensive reading on reading attitudes in a foreign language. *Reading in a Foreign Language*, 25, 248–263.
- Young, D. J. (1999). Reading anxiety scale (RAS). In D. J. Young (Ed.), *Affect in foreign language and second language learning: a practical guide to creating a low-anxiety classroom atmosphere* (p. 256). New York: McGraw Hill.
- Zhao, A., Guo, Y., & Dynia, J. (2013). Foreign language reading anxiety: Chinese as a foreign language in the United States. *The Modern Language Journal*, 97, 764–778.
<https://doi.org/10.1111/j.1540-4781.2013.12032.x>
- Zhou, J. (2017). Foreign language reading anxiety in a Chinese as a foreign language context. *Reading in a Foreign Language*, 29, 155–173.

Appendix

Means and Standard Deviations on the Anxiety Indexes

Questionnaire items	<i>N</i> = 57		<i>N</i> = 56	
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>
A) FLRAS (1–20)				
1. The hardest part of learning English is learning to read.	2.89	1.01	3.13	1.21
2. I would be happy just to learn to speak English rather than having to learn to read as well.	2.11	1.14	2.07	1.17
A-1) Grammar and Vocabulary (3–9)				
3. I get upset when I'm not sure whether I understand what I am reading in English.	3.51	1.10	3.93	0.97
4. I am nervous when I am reading a passage in English when I am not familiar with the topic.	3.30	1.16	3.57	1.14
5. I get upset whenever I encounter unknown grammar when reading English.	3.63	1.03	3.50	1.11
6. It bothers me to encounter words I can't pronounce while reading English.	3.23	1.04	3.18	1.19
7. You have to know so much about English history and culture in order to read English.	3.37	1.16	2.77	1.14
8. When reading English, I get nervous and confused when I don't understand every word.	3.21	1.22	3.30	1.20
9. When reading English, I often understand the words but still can't quite understand what the author is saying.	3.37	1.10	3.48	1.01
A-2) Language Distance (8, 10–15)				
10. I feel intimidated whenever I see a whole page of English in front of me.	2.46	1.20	2.59	1.32
11. I usually end up translating word by word when I'm reading English.	2.56	1.28	2.34	1.05
12. By the time you get past the unknown words in English, it's hard to remember what you're reading about.	3.60	1.03	3.84	1.12
13. I am worried about all the new words you have to learn in order to read English.	3.14	1.14	3.41	1.30
14. I don't mind reading to myself, but I feel very uncomfortable when I have to read English aloud.	2.46	1.25	2.27	1.10
15. English culture and ideas seem very foreign to me.	3.18	1.07	2.93	1.26

A-3) *Confidence/Enjoyment* (9, 16–20)

16. When I'm reading English, I get so confused I can't remember what I'm reading.	2.88	1.02	3.14	1.33
17. I am satisfied with the level of reading ability in English that I have achieved so far. ^R	4.21	1.03	4.07	1.08
18. I enjoy reading English. ^R	2.79	1.05	3.05	1.00
19. I feel confident when I am reading in English. ^R	3.63	1.10	3.54	1.01
20. Once you get used to it, reading English is not so difficult. ^R	2.98	1.06	2.82	1.11

B) **MBS** (21–24)

21. I become anxious when I have to read in English.	2.68	1.18	2.91	1.31
22. I fear having to read lengthy texts in English.	3.05	1.16	3.25	1.25
23. I fear not understanding the lengthy texts.	3.28	1.08	3.64	1.14
24. I become anxious when I have to answer the questions about what I have read in English.	2.91	1.12	3.09	1.31

C) **YAS** (25–28)

25. I feel anxious if I don't know all the words.	3.44	1.39	3.71	1.16
26. I sometimes feel anxious that I may not understand even if I read.	3.51	1.10	3.93	0.97
27. I feel anxious when I'm not sure whether I understood the texts.	3.28	1.33	3.34	1.08
28. I don't mind even if I cannot understand the contexts of reading material entirely. ^R	2.79	1.35	2.96	1.17

Note. ^R = reverse-scored items

About the Author

Hitoshi Mikami is a faculty member in the Department of English Language and Culture at Chubu University. His research interests include individual differences in language learning and language learning in study abroad contexts. Email: mikami_h@isc.chubu.ac.jp.