Analysis of Japanese EFL College Students' L1-L2 Reading Anxiety

Sachi Oshima
Chuo Gakuin University

Reference Data:

Oshima, S. (2026). Analysis of Japanese EFL college students' L1-L2 reading anxiety. In B. Lacy, M. Swanson, & H. Yoder (Eds.), *LanguageS: Learning, Teaching, Assessing – JALT 50 Years – Challenges and Perspectives*. JALT. https://doi.org/10.37546/JALTPCP2025-15

Since the introduction of the Foreign Language Reading Anxiety Scale (FLRAS; Saito et al., 1999), it has been widely used to assess learners' reading anxiety. However, when administered to Japanese college students, the FLRAS has issues, leading to the development of a new reading anxiety questionnaire (Oshima, 2024). Although previous researchers (e.g., Ghaith, 2020) have shown that reading anxiety negatively affects comprehension, there remains limited understanding of how Japanese students' reading anxiety differs between their first language (L1; Japanese) and second language (L2; English). In this study, 50 Japanese college students with low English proficiency completed the new questionnaire twice: for L1, and for L2. Results showed significantly higher anxiety in L2 reading. Three patterns were also identified: low L1 and high L2 anxiety; low L1 but varied L2 anxiety; and notable L1 anxiety among some students. These findings highlight the need to consider L1 reading anxiety alongside L2 anxiety.

Foreign Language Reading Anxiety Scale (FLRAS; Saitoほか, 1999)が発表されて以来、FLRASは学習者のリーディングに対する不安感（reading anxiety）測定に広く使用されてきた。しかし日本の大学生に使用する際に懸念点があることから、新たなアンケートを作成した（Oshima, 2024）。先行研究（Ghaith, 2020など）ではreading anxietyは読解に悪影響を及ぼすと指摘されているが、日本人大学生が第1言語（日本語）と第2言語（英語）を読む際の違いについてはあまり検証されていない。そこで本研究では、英語初級レベルの日本人大学生50名がアンケートに2度（1度目は日本語、2度目は英語の読解に対し）回答した。全体的に第2言語に対する高い不安感が示された。また、 (a) 第1言語への不安感は低く、第2言語への不安感が高い、(b) 第1言語への不安感は低く、第2言語への不安感は個人差がある、(c)第1言語に顕著な不安感を示す学生もいる、という3つのパターンも見られた。第2言語だけでなく第1言語に対するReading anxietyも考慮する必要性が示された。

Reading is one of the four fundamental skills in language learning—alongside listening, speaking, and writing—and plays an essential role in acquiring a foreign language (FL) or second language (L2). Among the variables affecting reading, Saito et al. (1999) proposed a construct, foreign language reading anxiety (FLRA), which is a skill-specific form of anxiety that FL learners experience during reading. They also developed the Foreign Language Reading Anxiety Scale (FLRAS) to measure this construct. Since its development, FLRAS has been widely used by researchers to measure learners' reading anxiety and its impact on reading performance (e.g., Chen et al., 2022; Ghaith, 2020; Lu & Liu, 2015; Zhao et al., 2025). However, its application to Japanese college students with lower English proficiency, who typically study English as a foreign language (EFL), requires careful consideration. According to Oshima (2024), key issues regarding FLRAS include inconsistent wording of anxiety-related adjectives in the questionnaire items, the presence of double-barreled items, and the inclusion of a midpoint on the five-point Likert scale. To address these issues, a new reading anxiety questionnaire (Oshima, 2024; see Appendix A) was created specifically for Japanese college students with lower English proficiency.

Recently, Japanese students' low language proficiency has not merely been a matter of learning L2 English. The latest comparative analytical survey conducted by the Ministry of Education, Sports, Science and Technology of Japan (MEXT, 2025) revealed the significant decline in academic performance of third-year junior high school students (typically aged 14 to 15) from fiscal year 2021 to 2024—not only in L2 English but also, notably, in their L1 Japanese. It is unsurprising therefore that students with low L2 English proficiency might find it difficult to read and comprehend English instruction. However, a decline in L1 Japanese competence can pose more serious obstacles, potentially hindering students' understanding of all subjects—even those taught in their L1. In response, some universities, including the researcher's own institution, have already introduced a compulsory L1 Japanese course for all first-year college students.

Literature Review

FLRA and Its Impacts on Reading Comprehension

Since the development of FLRAS (Saito et al., 1999), previous researchers have used this instrument to measure learners' FLRA and shared the notion that FLRA or L2 reading anxiety negatively affects reading comprehension (e.g., Ghaith, 2020; Lu & Liu, 2015; Saito et al., 1999; Zhao et al., 2025). For instance, Ghaith (2020), Lu and Liu (2015), and Zhao et al. (2025) all employed Saito et al's FLRAS to measure L2 English reading anxiety of their participants whose L1 was Arabic (Ghaith, 2020) or Chinese (Lu & Liu, 2015; Zhao et al., 2025).

Among these previous researchers, Ghaith (2020) examined the direct and indirect roles of FLRA on L1 Arabic EFL learners' reading comprehension (N = 103, aged 17 to 28). The results of path analyses suggested that FLRA negatively impacted reading comprehension ($\beta = -.31$, $p < .01$). In line with Ghaith (2020), Zhao et al. (2025) targeted Chinese EFL learners (N = 147, aged 17 to 20) and examined the roles of L1 reading comprehension, L2 vocabulary and syntactic knowledge, L2 reading strategy use, and L2 reading anxiety in L2 reading comprehension. Using hierarchical regression model analyses, Zhao et al. (2025) found that L2 reading anxiety significantly predicted L2 reading comprehension ($t = -2.35$, $p < .05$) and negatively affected learners across different L2 proficiency levels.

Research Gap and the Purpose of This Study

As described above, MEXT's (2025) survey comparing 3rd-year junior high school students' literacy skills between 2021 and 2024 revealed a significant decline not only in L2 English but also in L1 Japanese. The Japanese education system is generally called the 6-3-3 system, consisting of six years of elementary school, three years of junior high school, and three years of senior high school. The students who participated in MEXT's survey, conducted in 2024, are 2nd-year senior high school students as of 2026, meaning that they will enter university a few years later. This trend implies that L1 Japanese reading anxiety might be more common than previously assumed.

Nevertheless, compared with a larger number of studies on FLRA, as shown by Li's (2022) meta-analysis of 40 FLRA-related articles (involving 9,785 participants) published between 1998 and 2018, less attention has been paid to L1 reading anxiety and its effects. Piccolo et al. (2017) reviewed the negative influences of L1 reading anxiety on children's reading performance. They pointed out that "although reading anxiety has been largely studied in the context of second language acquisition (L2), there is a lack of new research on reading anxiety related to first language learning (L1)" (p. 537). Indeed, previous researchers examined L1-related variables; however, L1 reading anxiety has not been specifically examined. For instance, Sparks and Alamer (2023) showed that L1 achievement, assessed by several tests, including L1 vocabulary and L1 reading comprehension tests, had both direct and indirect effects on L2 reading anxiety. Similarly, the aforementioned Zhao et al. (2025) demonstrated that L1 reading comprehension was a significant predictor of L2 reading comprehension. Although both groups of researchers considered the influence of L1 reading performance on L2 reading, neither directly examined L1 reading anxiety itself.

Given the limited empirical attention to L1 reading anxiety, further investigation is needed. To partially address this research gap, the following research question will be examined: How similar or different is the reading anxiety of Japanese EFL college students when they read L1 Japanese and L2 English? As such, the purpose of this study was to provide empirical evidence by comparing L1 and L2 reading anxiety using Oshima's (2024) newly developed questionnaire.

Methodology

Participants

Fifty Japanese EFL college students with low English proficiency (aged 18 to 22) were asked to complete a newly developed questionnaire (Oshima, 2024) twice. All the students were non-English majors, and their English proficiency was estimated to be at the A1 level of the Common European Framework of Reference for Languages (CEFR; Council of Europe, 2026) based on their scores on the TOEIC Bridge® Listening & Reading Tests (Institute for International Business Communication, n.d.). They were selected through convenience sampling, given that they were in the researcher's intact English classes. All the students provided written informed consent to use their data, and this project was approved by the university's administrators.

All these students must take compulsory L1 Japanese courses once a week (90 minutes) in their first year, whereas they do not necessarily learn L2 English. They can choose one of the following FLs as a compulsory course: French, German, Chinese, Korean, or English. The students who participated in this study were those who had chosen English, and they took the university's compulsory English Reading and Grammar course once a week (90 minutes) in their first and second years.

Procedure

The newly developed reading anxiety questionnaire (Oshima, 2024; see Appendix A) was administered to the students in their L1, Japanese, during the lesson time (approximately 15 minutes) as a paper-based survey under the researcher's supervision in a classroom. The questionnaire consisted of two sections—Section I (Self-Focused Apprehension) and Section II (Task-Focused Apprehension). Each section contained 10 items rated on a six-point Likert scale (1 = *strongly disagree* to 6 = *strongly agree*).

In September 2023, students were asked to express their feelings when reading their L1 Japanese, and four weeks later (in October 2023), they were asked to do the same for their L2 English. The questionnaire items were identical, except for whether students were answering about L1 Japanese reading or L2 English reading. The four-week gap between the two questionnaires was strategically set to help prevent potential bias from the students' recall of their initial responses.

Data Analysis

The students' responses to each question were entered into a Microsoft Excel spreadsheet and saved as a comma-separated values (CSV) file for analysis, using statistical software R version 4.6.1 (R Core Team, 2026) with the WRS2 package (version 1.1-7; Mair & Wilcox, 2020). As described above, the questionnaire consisted of two sections containing 10 items. Moreover, the students answered the same questionnaire in the scenario of reading their L1 Japanese or L2 English. Therefore, each questionnaire item was labeled, such as JS1Q1 (Japanese reading anxiety questionnaire, Section 1, Question 1). This allowed each of the 20 item-level paired comparisons (10 items in Section 1 and 10 items in Section 2) to be implemented.

The questionnaire employed a six-point Likert scale, which is technically ordinal and thus typically analyzed by non-parametric tests. From this point of view, when comparing questionnaire items on L1 Japanese and L2 English reading anxiety, non-parametric Wilcoxon signed-rank tests can be used because the data are ordinal. On the other hand, previous researchers have also supported treating Likert-scale responses as approximately interval and using more robust parametric tests (Brown, 2016; Huh & Gim, 2025). For instance, Brown (2016) stated that "whether Likert items are interval or ordinal is irrelevant in using Likert scale data, which can be taken to be interval" (p. 104). Huh and Gim (2025) suggested that the Likert scale data can be treated as interval data for practical research purposes, demonstrating that analyzing Likert data with parametric methods does not compromise statistical validity, especially when Likert scales include five or more points. Based on the six-point scale used in this study, the sample size ($N = 50$), and the aggregated nature of the scores, treating the data as approximately interval can be justified. For this reason, the use of parametric paired-samples *t*-tests was initially considered.

However, the assumption of normality, which is required for conducting parametric paired-samples *t*-tests, was not met in the present dataset, as indicated by statistically significant Shapiro-Wilk test results for 19 out of 20 items ($p < .05$). Due to this violation of the assumption, two analyses were selected for all 20 comparisons: a bootstrap-based Yuen's test and a non-parametric Wilcoxon signed-rank test. The command file is provided in Appendix B.

First, as the primary analysis, a bootstrap-based Yuen's test on 20% trimmed means was adopted, following the robust framework described by Field and Wilcox (2017). The WRS2 package (version 1.1-7) was used as the basis for the trimmed-mean test statistics (Mair & Wilcox, 2020). According to the WRS2 Reference Manual (Mair & Wilcox, 2025, pp. 42–45), the *yuend()* function allows paired-samples robust *t*-tests, providing values such as the test statistic (*t*-statistic), *p*-value, confidence interval (CI), and trimmed mean difference for paired samples. However, the bootstrap implementation *yuenbt()* is available only for independent samples and is therefore not applicable for paired-samples comparison in this study (L1 Japanese vs. L2 English). Accordingly, bootstrap percentile CIs were computed separately for the paired difference scores using the *boot* package (version 1.3-32) in R. The bootstrap procedure used 2,000 resamples, in line with the recommendation that 1,000 or 2,000 bootstrap samples are commonly used in practice (Field & Wilcox, 2017). All bootstrap-based procedures used a fixed random seed (123) for reproducibility. Second, as a complementary analysis, non-parametric Wilcoxon signed-rank test was also conducted. For both Yuen's test and the Wilcoxon signed-rank test, Holm's sequential procedure was applied across all 20 comparisons to control the familywise error rate (Holm, 1979). This method was chosen over the standard Bonferroni correction because it is suggested as more powerful while maintaining the same Type I error control (Aickin & Gensler, 1996; Holm, 1979).

Results

Overall Differences between L1 and L2 Reading Anxiety

The bootstrap-based Yuen's test revealed statistically significant differences between students' L1 Japanese and L2 English reading anxiety across all 20 questionnaire items (Table 1). All difference scores were computed as L1 Japanese–L2 English; negative values indicate lower L1 reading anxiety scores compared with L2 scores.

Table 1
Results of the Bootstrap-based Yuen's Test with Holm Correction

Item	L1 *M*	L2 *M*	Trimmed Diff	Boot CI *LL*	Boot CI *UL*	*t*	*df*	*p*	*p* (*Holm*)	ξ
S1Q1	2.17	3.90	−1.73	−2.00	−1.00	−6.56	29	< .001	< .001	0.87
S1Q2	1.70	3.40	−1.70	−1.90	−0.97	−5.54	29	< .001	< .001	0.70
S1Q3	1.63	3.23	−1.60	−1.87	−1.00	−5.37	29	< .001	< .001	0.68
S1Q4	2.17	3.50	−1.33	−1.50	−0.70	−5.45	29	< .001	< .001	0.69
S1Q5	2.73	4.40	−1.67	−2.07	−1.00	−6.14	29	< .001	< .001	0.65
S1Q6	1.97	2.97	−1.00	−1.27	−0.43	−4.26	29	< .001	= .001	0.43
S1Q7	2.40	4.07	−1.67	−1.77	−0.93	−8.19	29	< .001	< .001	0.84
S1Q8	2.47	4.13	−1.67	−1.90	−0.97	−8.43	29	< .001	< .001	0.66
S1Q9	2.13	3.70	−1.57	−1.63	−0.70	−5.98	29	< .001	< .001	0.62
S1Q10	1.97	3.33	−1.37	−1.67	−0.73	−4.39	29	< .001	< .001	0.71
S2Q1	2.47	3.87	−1.40	−1.63	−0.73	−5.67	29	< .001	< .001	0.71
S2Q2	3.07	4.07	−1.00	−1.27	−0.30	−4.01	29	< .001	.002	0.41
S2Q3	2.93	3.70	−0.77	−1.10	−0.30	−3.22	29	.003	.007	0.39
S2Q4	2.87	3.97	−1.10	−1.30	−0.50	−5.08	29	< .001	< .001	0.45
S2Q5	2.20	3.23	−1.03	−1.13	−0.33	−3.62	29	.001	.004	0.54
S2Q6	2.43	3.03	−0.60	−0.83	−0.17	−2.55	29	.016	.016	0.32
S2Q7	2.70	4.03	−1.33	−1.50	−0.60	−5.56	29	< .001	< .001	0.49
S2Q8	2.70	3.53	−0.83	−1.03	−0.27	−3.33	29	.002	.007	0.42
S2Q9	2.87	4.27	−1.40	−1.63	−0.60	−5.56	29	< .001	< .001	0.53
S2Q10	2.63	3.73	−1.10	−1.30	−0.47	−4.55	29	< .001	< .001	0.57

Note. $N = 50$. S= Section; Q = Question; *M* = Mean; Trimmed Diff = Trimmed Mean Difference; Boot CI = Bootstrap Confidence Interval; *LL* = Lower Limit; *UL* = Upper Limit; *p* (Holm) = Holm-adjusted *p*-values. Statistical tests were conducted using Yuen's 20% trimmed mean paired-samples *t*-test. Effect sizes are reported as ξ (xi). See Appendix A for questionnaire items and Appendix B for R command.

As Table 1 shows, all 20 item-level paired comparisons showed statistically significant differences after Holm correction (all adjusted $p < .05$); L1 Japanese reading anxiety scores were consistently lower than L2 English reading anxiety scores. Negative trimmed mean differences (ranging from −1.73 to −0.60) showed that, across all 20 items, the students reported lower anxiety when reading in L1 Japanese compared to L2 English. Effect sizes (ξ) ranged from 0.32 to 0.87. According to Mar and Wilcox's (2020) criteria, suggesting that "values of $\xi = 0.10$, 0.30, and 0.50 correspond to small, medium, and large effect sizes" (p. 469), 13 out of 20 items showed large effects ($\xi \geq 0.50$) and the remaining seven showed medium effects ($\xi \geq .30$).

As a complementary non-parametric analysis, the Wilcoxon signed-rank test was also conducted, showing that all 20 item-level paired comparisons were statistically significant. The results were consistent with those of the bootstrap-based Yuen's test (see Appendix C).

Three Patterns of L1–L2 Differences

Although the students overall showed lower anxiety toward L1 Japanese reading than toward L2 English reading (Table 1), further analysis of students' responses to individual questionnaire items revealed three distinct patterns as shown below:

Pattern 1: Low L1 Anxiety vs. High L2 Anxiety

Pattern 1 aligns with the overall tendencies described above; more specifically, the students reported lower anxiety when reading L1 Japanese but substantially higher anxiety when reading L2 English. Figure 1 shows that the skewness of each histogram was in the opposite direction depending on whether students read Japanese or English.

This pattern was especially evident in items related to *inferencing* or *guessing* the meaning of unfamiliar vocabulary or parts that the students found hard to understand, as shown in Section 1, Questions 1, 7, and 8:

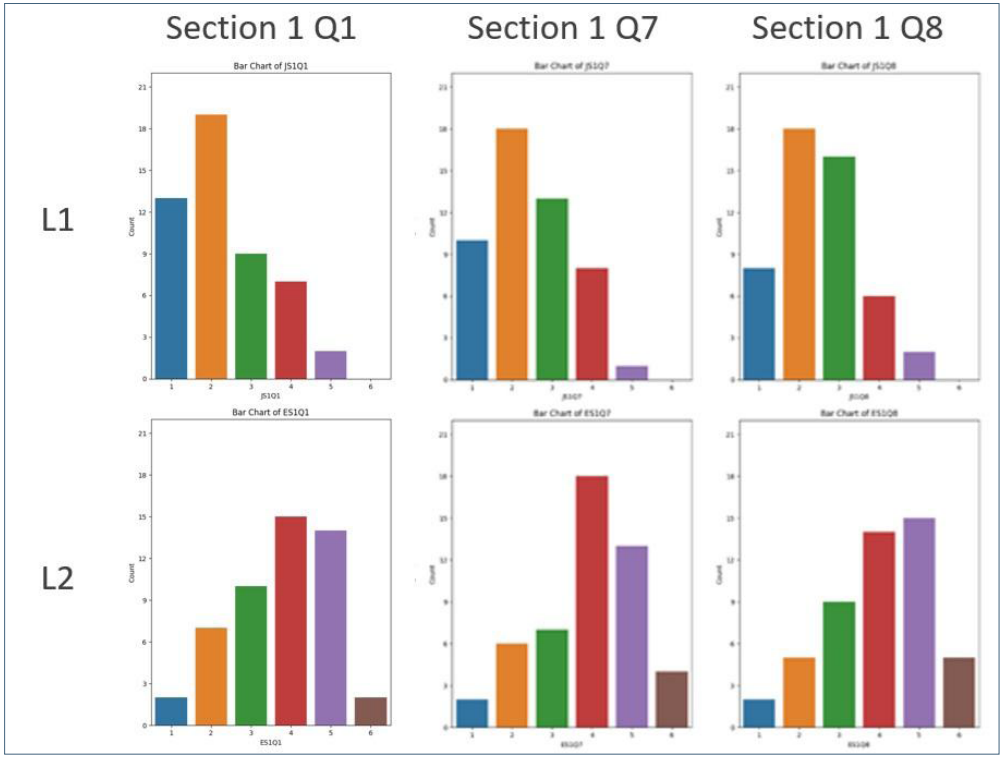
S1Q1: When reading Japanese/English, it is difficult to guess the meaning of parts I cannot understand.

S1Q7: I get confused when reading new information in Japanese/English.

S1Q8: When reading Japanese/English, it is difficult to guess the meaning of new words.

Figure 1

Students' Responses to Section 1 Questions 1, 7, and 8



Note. From left to right: 1 = Strongly disagree, 2 = Disagree, 3 = Slightly disagree, 4 = Slightly agree, 5 = Agree, 6 = Strongly agree.

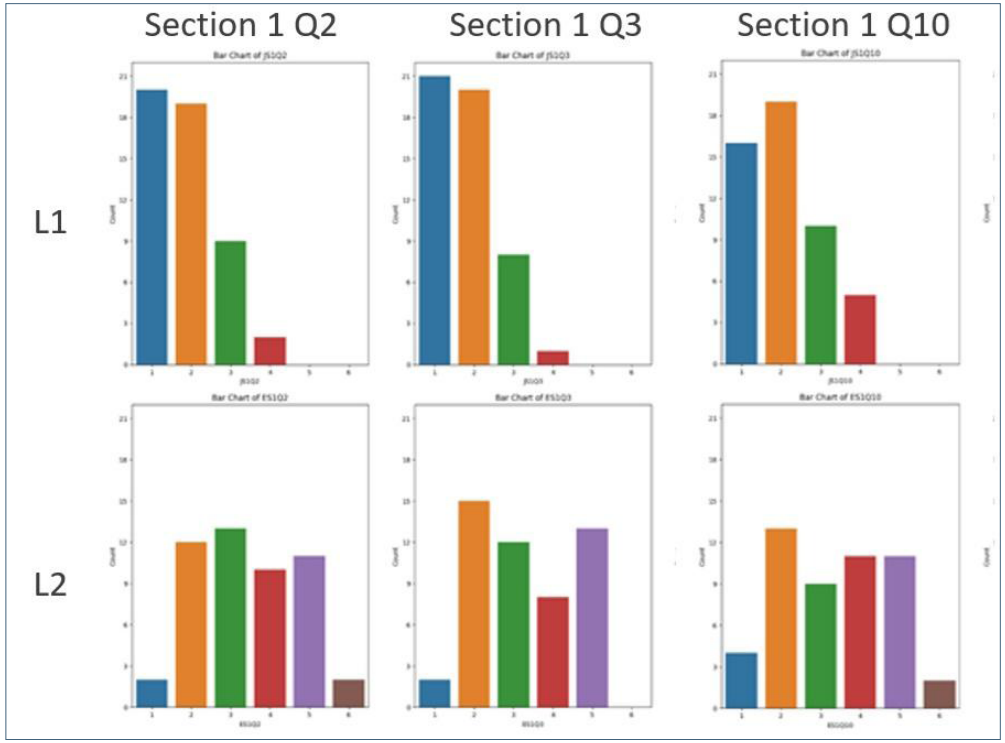
Pattern 2: Low L1 Anxiety with Varied L2 Anxiety

The students showed consistently lower anxiety in L1 Japanese reading, whereas their L2 English anxiety varied depending on each student. This pattern suggests that, although individual differences were observed, most students felt comfortable or found each reading task less cognitively demanding in their L1. On the other hand, for low English proficiency students, even relatively simple tasks, such as textbook instruction (Section 1 Question 2), short passages (Section 1 Question 3), or texts that include already known vocabulary (Section 1 Question 10), led them to have higher L2 English reading anxiety (Figure 2).

- S1Q2: It is difficult to understand instructions written in Japanese/English in my Japanese/English textbooks.
- S1Q3: I often feel confused when reading short passages in Japanese/English.
- S1Q10: I often do not understand what Japanese/English texts say even when I know the words.

Figure 2

Students' Responses to Section 1 Questions 2, 3, and 10



Note. From left to right: 1 = Strongly disagree, 2 = Disagree, 3 = Slightly disagree, 4 = Slightly agree, 5 = Agree, 6 = Strongly agree.

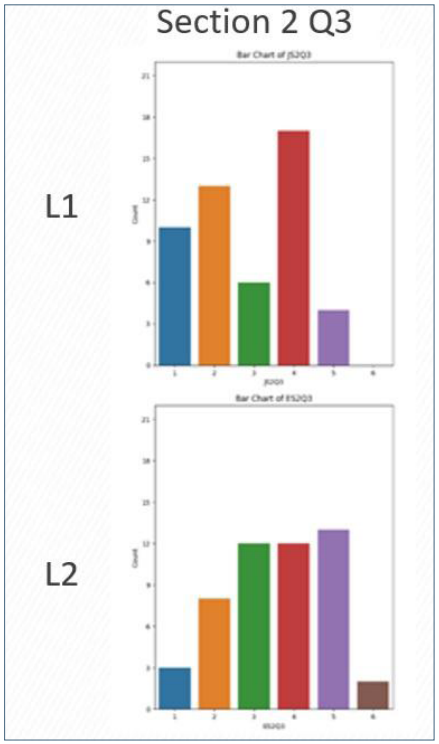
Pattern 3: High L1 Anxiety Shown among Some Students

Some students exhibited notable higher anxiety even when reading in their L1. These students reported that they would feel anxiety toward L1 texts without topic familiarity. On the other hand, their L2 English anxiety levels varied (Figure 3).

S2Q3: When reading Japanese/English, I feel worried when I'm not familiar with this topic.

Figure 3

Students' Responses to Section 2 Question 3



Note. From left to right: 1 = Strongly disagree, 2 = Disagree, 3 = Slightly disagree, 4 = Slightly agree, 5 = Agree, 6 = Strongly agree.

Other Variable to Note: Influence of Passage Length

In addition to the three patterns observed, the students' responses to the questionnaire items on passage length showed that longer texts increased anxiety in both L1 and L2 (Section 1, Questions 3 and 5). The perception of length varies depending on each individual; therefore, when students answered these two items, the length of passages was defined referring to the students' English reading textbook used in the lessons: one-paragraph long (approximately 75-word to 100-word long) as a *short* passage, and one-chapter long (approximately 500-word long) as a *long* passage. For Japanese passages, a *short* passage was about 200 characters and a *long* passage was about 1,000 characters, approximately twice the English word count. Although this 1:2 (English word: Japanese character) ratio does not fully account for factors such as lexical difficulty and sentence complexity, this ratio seems to be a reasonable basis because it has widely been recognized as a convention in the English-Japanese translation industry (Qualtranscribe, 2026; TBSJ, 2022). Kano et al. (2024) also supported this ratio, suggesting that the number of Japanese characters was approximately twice the number of English words upon examining the International Conference on Spoken Language Translation (IWSLT) English-Japanese training dataset.

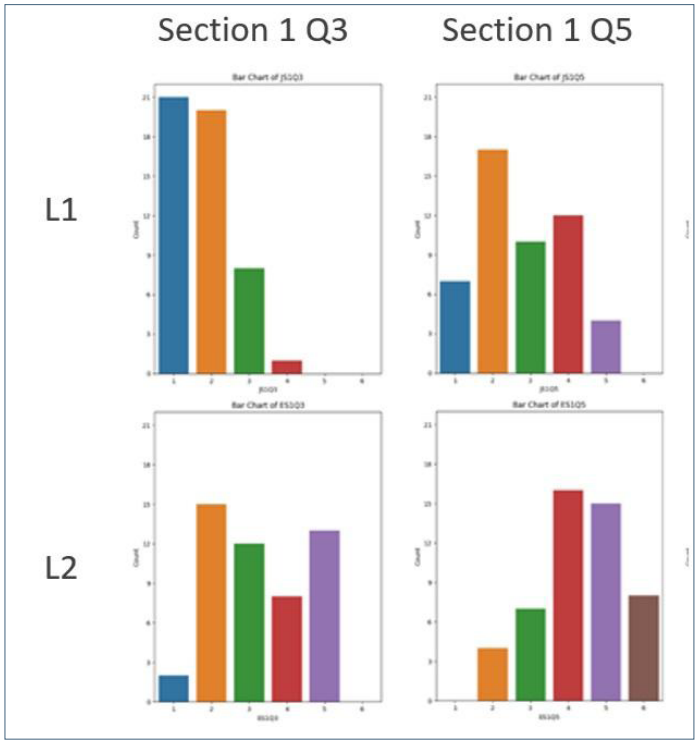
S1Q3: I often feel confused when reading short passages in Japanese/English.

S1Q5: I often feel confused when reading long passages in Japanese/English.

As Figure 4 shows, in L2 English reading, the students' reading anxiety remained similar, but slightly elevating when they were to read long passages. On the other hand, the students' reading anxiety was heightened with long passages in L1 Japanese reading. This indicates that text length is a cross-linguistic factor influencing reading anxiety, even among L1 speakers. Even in their L1, the students might struggle with sustained reading; in other words, the longer the passage is, the more anxiety students experience regardless of whether they read in L1 or L2.

Figure 4

Students' Responses to Section 1, Questions 3 and 5



Note. From left to right: 1 = Strongly disagree, 2 = Disagree, 3 = Slightly disagree, 4 = Slightly agree, 5 = Agree, 6 = Strongly agree.

Discussion

The results of this study demonstrated statistically significant differences in reading anxiety between Japanese low-English-proficiency college students' L1 Japanese and L2 English. Consistent with previous researchers on FLRA, demonstrating that L2 reading anxiety is a robust predictor of L2 reading difficulty (e.g., Ghaith, 2020; Lu & Liu, 2015; Saito et al., 1999; Zhao et al., 2025), low L2 proficiency students reported substantially higher anxiety when reading L2 English. This finding supports the idea that L2 reading puts cognitive and linguistic demands on learners, especially those with lower L2 English proficiency, which can heighten their L2 reading anxiety.

Three distinct patterns of L1–L2 differences were shown, offering a more detailed understanding of how reading anxiety emerges across languages. The first pattern—low L1 anxiety with high L2 anxiety—reflects the typical feature of low L2 proficiency learners, whose limited vocabulary and grammatical knowledge make their L2 English reading challenging. The second pattern, in which students showed low anxiety in L1 Japanese but varied levels in L2 English, suggests that relatively cognitively less demanding tasks can lead to strong anxiety for some students in L2 reading. The third pattern, where some students exhibited high anxiety even in their L1 depending on the topic familiarity, supports the notion that reading anxiety is not exclusively a FL phenomenon (Piccolo et al., 2017). Given recent reports of declining L1 literacy among Japanese students (MEXT, 2025), the presence of L1 reading anxiety among some college students is unsurprising, and further investigation is expected.

Additionally, the students reported increased anxiety when reading longer texts in both L1 Japanese and L2 English. This finding underscores the importance of task design and careful scaffolding of reading materials in both L1 and L2 contexts. For teachers, it seems important to address not only student L2 reading challenges but also potential L1 reading difficulties. Supporting students' general reading confidence may benefit reading in both languages, particularly for those with reading anxiety even in their L1.

In summary, the findings suggest that reading anxiety among Japanese college students with lower L2 proficiency is influenced by both language-related and individual factors. Although higher L2 reading anxiety has already been expected based on previous literature (e.g., Ghaith, 2020; Lu & Liu, 2015; Saito et al., 1999; Zhao et al., 2025), the presence of L1 reading anxiety among some students highlights the need to consider reading more holistically, regardless of L1 or L2. Supporting students' general reading confidence, rather than focusing solely on L2 reading skills, may have positive effects across languages.

Conclusion

In this study, the reading anxiety of Japanese EFL college students with lower English proficiency was examined by comparing reading in their L1 Japanese and L2 English. The results revealed significantly higher anxiety in L2 English reading across all 20 questionnaire items, which was not surprising in light of previous literature (e.g., Ghaith, 2020; Lu & Liu, 2015; Saito et al., 1999; Zhao et al., 2025). At the same time,

three patterns found in the students' responses suggest that L1 reading anxiety and L2 reading anxiety are not uniform but are affected by various factors, including linguistic difficulties, text length, and topic familiarity.

Although these findings can partially contribute to the limited research on L1 reading anxiety, several limitations should be acknowledged when interpreting the results. First, the sample size was small ($N = 50$), and the participants were recruited through convenience sampling from a single institution, which limits the generalizability of results. Second, among the 20 questionnaire items, one item (Section 1 Question 6: When reading Japanese/English, I often feel confused even when I translate word by word.) was not appropriate for the L1 Japanese reading context because it is unlikely for the students to translate their L1 Japanese to other languages when reading L1. Even so, the robust analysis confirmed a significant difference for this item, suggesting that the overall pattern of L1-L2 differences is not driven by any single item. Third, although the newly developed questionnaire (Oshima, 2024) was tailored to Japanese college students with lower L2 English proficiency, its applicability to other populations has not yet been validated.

Despite these limitations, a key finding of this study is that the students with low L2 proficiency reported anxiety even when reading in their L1. This suggests that reading anxiety does not solely concern L2. This finding aligns with recent concerns (MEXT, 2025) about a significant decline in L1 literacy among Japanese students. Accordingly, this study can contribute to an emerging area of inquiry by offering initial evidence from a Japanese university context. Future research with larger and more diverse samples is needed to further explore how students' reading anxiety, both in their L1 and L2, evolves and how it can be more effectively addressed in educational settings.

Acknowledgements

I would like to express my deepest gratitude to David Beglar for his invaluable guidance and encouragement during the development of the reading anxiety questionnaire, which forms the foundation of this study. I also appreciate the thoughtful feedback provided by all those who contributed helpful comments on the data and earlier drafts of the manuscript.

Bio Data

Sachi Oshima is an associate professor in the Faculty of Law at Chuo Gakuin University, Chiba, Japan. She began her career as a senior high school English teacher and later worked for the Japan Foundation as a chief officer in charge of intercultural exchange projects. Since 2017, she has been teaching English at the university level. Her research interests include TESOL and intercultural exchange. <oshima.s@mc.cgu.ac.jp>

References

Aickin, M., & Gensler, H. (1996). Adjusting for multiple testing when reporting research results: The Bonferroni vs Holm methods. *American Journal of Public Health*, *86*(5), 726–728. https://doi.org/10.2105/AJPH.86.5.726

Brown, J. D. (2016). *Statistics Corner: Questions and answers about language testing statistics*. JALT Testing and Evaluation Special Interest Group.

Chen, S., Du, H., Wang, S., & Yang, L. (2022). Understanding EFL reading anxiety in relation to learning motivation, attitudes and strategies for Chinese and Spanish undergraduates. *System*, *108*, Article 102842. https://doi.org/10.1016/j.system.2022.102842

Council of Europe. (2026). *The CEFR levels*. Retrieved August 24, 2026, from https://www.coe.int/en/web/common-european-framework-reference-languages/level-descriptions

Field, A. P., & Wilcox, R. R. (2017). Robust statistical methods: A primer for clinical psychology and experimental psychopathology researchers. *Behaviour Research and Therapy*, *98*, 19–38. https://doi.org/10.1016/j.brat.2017.05.013

Ghaith, G. M. (2020). Foreign language reading anxiety and metacognitive strategies in undergraduates' reading comprehension. *Issues in Educational Research*, *30*(4), 1310–1328. http://www.iier.org.au/iier30/ghaith.pdf

Huh, I., & Gim, J. (2025). Exploration of Likert scale in terms of continuous variable with parametric statistical methods. *BMC Medical Research Methodology*, *25*(1), Article 218. https://doi.org/10.1186/s12874-025-02668-1

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, *6*(2), 65–70. https://www.jstor.org/stable/4615733

Institute for International Business Communication. (n.d.). *TOEIC Bridge® Listening & Reading Tests*. Retrieved August 24, 2026, from https://www.iibc-global.org/english/toeic/test/bridge.html

Kano, Y., Sudoh, K., & Nakamura, S. (2024). Average token delay: A duration-aware latency metric for simultaneous translation. *Journal of Natural Language Processing*, *31*(3), 1049–1075. https://doi.org/10.5715/jnlp.31.1049

Li, R. (2022). Foreign language reading anxiety and its correlates: A meta-analysis. *Reading and Writing*, *35*(4), 995–1018. https://doi.org/10.1007/s11145-021-10213-x

Loerts, H., Lowie, W., & Seton, B. (2020). *Essential statistics for applied linguistics: Using R or JASP* (2nd ed.). Red Globe Press.

Lu, Z., & Liu, M. (2015). An investigation of Chinese university EFL learner's foreign language reading anxiety, reading strategy use and reading comprehension performance. *Studies in Second Language Learning and Teaching*, *5*(1), 65–85. https://doi.org/10.14746/ssllt.2015.5.1.4

Mair, P., & Wilcox, R. (2020). Robust statistical methods in R using the WRS2 package. *Behavior Research Methods*, *52*, 464–488. https://doi.org/10.3758/s13428-019-01246-w

Mair, P., & Wilcox, R. (2025). *WRS2: A collection of robust statistical methods* (R package version 1.1-7) [Computer software manual]. CRAN. https://cran.r-project.org/web/packages/WRS2/WRS2.pdf

Ministry of Education, Culture, Sports, Science and Technology of Japan (MEXT). (2025). *Reiwa 7nendo zenkoku gakuryoku gakushu jokyo chosa no houkokusho, shukei kekka ni tsuite* [Report and aggregate results of the fiscal 2025 national survey for academic achievement]. https://www.mext.go.jp/a_menu/shotou/gakuryoku-chousa/sonota/1419141_00007.htm

Oshima, S. (2024). Developing the reading anxiety questionnaire for Japanese beginner-level college students. In M. Campos, A. MacKenzie, E. Vasquez, D. Iske, A. Minami, & S. Hutchinson (Eds.), *2023 PanSIG Journal* (pp. 125–132). Japan Association for Language Teaching (JALT). https://pansig.org/publications/2023/2023_PanSIG_Journal.pdf

Piccolo, L. R., Giacomoni, C. H., Julio-Costa, A., Oliveira, S., Zbornik, J., Haase, V. G., & Salles, J. F. (2017). Reading anxiety in L1: Reviewing the concept. *Early Childhood Education Journal*, *45*(4), 537–543. https://doi.org/10.1007/s10643-016-0822-x

Qualtranscribe. (2026). *From yen to words: A clear look at Japanese translation pricing*. https://www.qualtranscribe.com/blog/japanese-translation-pricing-guide

R Core Team. (2026). *R: A language and environment for statistical computing* (Version 4.6.1) [Computer software]. R Foundation for Statistical Computing. https://www.R-project.org/

Saito, Y., Horwitz, E. K, & Garza, T J. (1999). Foreign language reading anxiety. *The Modern Language Journal*, *83*(2), 202–218. https://doi.org/10.1037/t54908-000

Sparks, R. L., & Alamer, A. (2023). How does first language achievement impact second language reading anxiety? Exploration of mediator variables. *Reading and Writing*, *36*, 2667–2692. https://doi.org/10.1007/s11145-022-10410-2

TBSJ. (2022). *Frequently asked questions*. https://secure.tbsj.jp/Home/FAQ

Zhao, A., Zhang, X., Guo, Y., Dinnesen, M. S., Sun, S., & Xu, Y. (2025). Exploring the roles of L1 reading, L2 proficiency, strategy use and anxiety in L2 reading comprehension. *Journal of Research in Reading, 48*(1), 100–116. https://doi.org/10.1111/1467-9817.12478

Appendix A

Reading Anxiety Questionnaire (Oshima, 2024)

- 1 = Strongly disagree
- 2 = Disagree
- 3 = Slightly disagree
- 4 = Slightly agree
- 5 = Agree
- 6 = Strongly agree

Section I: Self-Focused Apprehension

1. When reading Japanese/English, it is difficult to guess the meaning of parts I cannot understand.
2. It is difficult to understand instructions written in Japanese/English in my English textbooks.
3. I often feel confused when reading short passages in Japanese/English.
4. I often cannot remember what I have read in Japanese/English.
5. I often feel confused when reading long passages in Japanese/English.
6. When reading Japanese/English, I often feel confused even when I translate word by word.
7. I get confused when reading new information in Japanese/English.
8. When reading Japanese/English, it is difficult to guess the meaning of new words.
9. Japanese/English spelling is confusing.
10. Even when I know the Japanese/English words, I often cannot understand what the Japanese/English sentence intends to say.

Section II: Task-Focused Apprehension

1. When reading Japanese/English, I feel worried when I come across one or two unfamiliar words.
2. When reading Japanese/English, I feel worried if I only read a passage once.
3. When reading Japanese/English, I feel worried when I'm not familiar with the topic.

4. When reading Japanese/English, I worry that I might have missed important information.
5. When reading Japanese/English, I feel worried when I don't understand every word.
6. When reading Japanese/English, I worry I might have inadequate knowledge about the topic.
7. I feel worried when I have little time to think about what I have read in Japanese/English.
8. I feel worried when I cannot read Japanese/English at a comfortable pace.
9. I feel worried when I'm not sure whether I have understood what reading texts say.
10. When reading Japanese/English, I worry that I may be missing key words.

Appendix B

R Scripts Used in This Study

The following R scripts were used to conduct all statistical analyses reported in this study. All analyses were conducted in R version 4.6.1 (R Core Team, 2026) using the WRS2 package (version 1.1-7; Mair & Wilcox, 2020) and the boot package (version 1.3-32). A fixed random seed (123) was applied to all bootstrap-based procedures to ensure reproducibility. Differences are reported as JS–ES throughout.

1. Normality Test (Shapiro-Wilk)

```
# Load data
df <- read.csv("L1-L2 relationship.csv", stringsAsFactors = FALSE)
df <- df[!is.na(df$Student), ]
cat("N participants:", nrow(df), "\n\n")

# Define the 20 pairs
pairs <- list(
  c("JS1Q1","ES1Q1"), c("JS1Q2","ES1Q2"), c("JS1Q3","ES1Q3"),
  c("JS1Q4","ES1Q4"), c("JS1Q5","ES1Q5"), c("JS1Q6","ES1Q6"),
  c("JS1Q7","ES1Q7"), c("JS1Q8","ES1Q8"), c("JS1Q9","ES1Q9"),
  c("JS1Q10","ES1Q10"),
  c("JS2Q1","ES2Q1"), c("JS2Q2","ES2Q2"), c("JS2Q3","ES2Q3"),
  c("JS2Q4","ES2Q4"), c("JS2Q5","ES2Q5"), c("JS2Q6","ES2Q6"),
  c("JS2Q7","ES2Q7"), c("JS2Q8","ES2Q8"), c("JS2Q9","ES2Q9"),
  c("JS2Q10","ES2Q10")
)

# Run Shapiro-Wilk test for each pair
normality_results <- data.frame(
  Pair = character(20), Section = character(20), Item = character(20),
  N = integer(20), W = numeric(20), p_value = numeric(20),
  Normality = character(20), stringsAsFactors = FALSE
)

for (i in seq_along(pairs)) {
  js_col <- pairs[[i]][1]
  es_col <- pairs[[i]][2]
  diff_scores <- df[[js_col]] - df[[es_col]]  # JS - ES

  sw <- shapiro.test(diff_scores)

  normality_results$Pair[i]      <- paste0(js_col, " vs ", es_col)
  normality_results$Section[i]   <- ifelse(grepl("JS1", js_col), "Section1",
    "Section2")
  normality_results$Item[i]      <- sub(".*Q", "Q", js_col)
  normality_results$N[i]         <- length(diff_scores)
  normality_results$W[i]         <- round(sw$statistic, 4)
  normality_results$p_value[i]   <- round(sw$p.value, 4)
  normality_results$Normality[i] <- ifelse(sw$p.value < 0.05, "Violated",
    "Met")
}

print(normality_results, row.names = FALSE)
write.csv(normality_results, "normality_results.csv", row.names = FALSE)
```

2. Bootstrap-based Yuen's Paired-Samples Test (20% Trimmed Means, 2,000 Resamples) with Holm Correction

```
library(WRS2)
library(boot)

# Load data
df <- read.csv("L1-L2 relationship.csv", stringsAsFactors = FALSE)
df <- df[!is.na(df$Student), ]
cat("N participants:", nrow(df), "\n\n")

# Define the 20 pairs
pairs <- list(
  c("JS1Q1","ES1Q1"), c("JS1Q2","ES1Q2"), c("JS1Q3","ES1Q3"),
  c("JS1Q4","ES1Q4"), c("JS1Q5","ES1Q5"), c("JS1Q6","ES1Q6"),
  c("JS1Q7","ES1Q7"), c("JS1Q8","ES1Q8"), c("JS1Q9","ES1Q9"),
  c("JS1Q10","ES1Q10"),
  c("JS2Q1","ES2Q1"), c("JS2Q2","ES2Q2"), c("JS2Q3","ES2Q3"),
  c("JS2Q4","ES2Q4"), c("JS2Q5","ES2Q5"), c("JS2Q6","ES2Q6"),
  c("JS2Q7","ES2Q7"), c("JS2Q8","ES2Q8"), c("JS2Q9","ES2Q9"),
  c("JS2Q10","ES2Q10")
)

# Bootstrap statistic: 20% trimmed mean of difference scores (JS - ES)
trim_mean_diff <- function(data, indices, tr = 0.2) {
  d <- data[indices]
  return(mean(d, trim = tr))
}

# Combined results data frame
combined_results <- data.frame(
  Item = character(20), Section = character(20),
  Trim_mean_JS = numeric(20), Trim_mean_ES = numeric(20),
  Trimmed_Diff = numeric(20),
  Boot_CI_Lower = numeric(20), Boot_CI_Upper = numeric(20),
  t_statistic = numeric(20), df = integer(20),
  p_value_raw = numeric(20), p_value_holm = numeric(20),
  Effect_size = numeric(20), stringsAsFactors = FALSE
)

set.seed(123)

for (i in seq_along(pairs)) {
  js_col <- pairs[[i]][1]
  es_col <- pairs[[i]][2]
  js <- df[[js_col]]
  es <- df[[es_col]]
  diff_scores <- js - es  # JS - ES

  # Yuen's test: yuend(x, y) computes x - y, so pass (js, es)
  yuen_out <- yuend(js, es, tr = 0.2)

  # Extract effect size from printed output
  yuen_txt <- capture.output(print(yuen_out))
  es_line <- grep("Explanatory", yuen_txt, value = TRUE)
  es_val <- if (length(es_line) >= 1) {
    nums <- regmatches(es_line[1], gregexpr("[0-9]+\\.[0-9]+|[0-9]+", es_
     line[1]))
    suppressWarnings(as.numeric(nums[[1]][1]))
  } else NA_real_

  # Bootstrap CI (2,000 resamples)
  boot_out <- boot(diff_scores, statistic = trim_mean_diff, R = 2000, tr =
     0.2)
  ci <- boot.ci(boot_out, type = "perc")$percent[4:5]

  combined_results$Item[i]          <- sub(".*Q", "Q", js_col)
  combined_results$Section[i]       <- ifelse(grepl("JS1", js_col),
     "Section1", "Section2")
  combined_results$Trim_mean_JS[i]  <- round(mean(js, trim = 0.2), 4)
```

```
  combined_results$Trim_mean_ES[i]  <- round(mean(es, trim = 0.2), 4)
  combined_results$Trimmed_Diff[i]  <- round(yuen_out$dif, 4)
  combined_results$Boot_CI_Lower[i] <- round(ci[1], 4)
  combined_results$Boot_CI_Upper[i] <- round(ci[2], 4)
  combined_results$t_statistic[i]   <- round(yuen_out$test, 4)
  combined_results$df[i]            <- yuen_out$df
  combined_results$p_value_raw[i]   <- yuen_out$p.value
  combined_results$Effect_size[i]   <- es_val

  cat(sprintf("Item %d/20 done: %s\n", i, js_col))
}

# Holm correction
combined_results$p_value_holm <- p.adjust(combined_results$p_value_raw,
     method = "holm")
combined_results$Sig <- ifelse(combined_results$p_value_holm < 0.05, "Yes",
     "No")

print(combined_results, row.names = FALSE)
write.csv(combined_results, "combined_yuen_bootstrap_holm_results_JS-ES.csv",
     row.names = FALSE)
```

3. Wilcoxon Signed-Rank Test with Holm Correction

```
# Load data
df <- read.csv("L1-L2 relationship.csv", stringsAsFactors = FALSE)
df <- df[!is.na(df$Student), ]
cat("N participants:", nrow(df), "\n\n")

# Define the 20 pairs
pairs <- list(
  c("JS1Q1","ES1Q1"), c("JS1Q2","ES1Q2"), c("JS1Q3","ES1Q3"),
  c("JS1Q4","ES1Q4"), c("JS1Q5","ES1Q5"), c("JS1Q6","ES1Q6"),
  c("JS1Q7","ES1Q7"), c("JS1Q8","ES1Q8"), c("JS1Q9","ES1Q9"),
  c("JS1Q10","ES1Q10"),
  c("JS2Q1","ES2Q1"), c("JS2Q2","ES2Q2"), c("JS2Q3","ES2Q3"),
  c("JS2Q4","ES2Q4"), c("JS2Q5","ES2Q5"), c("JS2Q6","ES2Q6"),
  c("JS2Q7","ES2Q7"), c("JS2Q8","ES2Q8"), c("JS2Q9","ES2Q9"),
  c("JS2Q10","ES2Q10")
)

# Run Wilcoxon signed-rank test for each pair (JS - ES direction)
n <- nrow(df)
wilcox_results <- data.frame(
  Item = character(20), Section = character(20),
  Median_JS = numeric(20), Median_ES = numeric(20),
  W_statistic = numeric(20), Z_statistic = numeric(20),
  p_value_raw = numeric(20), Effect_r = numeric(20),
  p_value_holm = numeric(20), Sig = character(20),
  stringsAsFactors = FALSE
)

for (i in seq_along(pairs)) {
  js_col <- pairs[[i]][1]
  es_col <- pairs[[i]][2]
  js <- df[[js_col]]
  es <- df[[es_col]]

  # wilcox.test(x, y) computes x - y, so pass (js, es) for JS - ES
  out <- wilcox.test(js, es, paired = TRUE, exact = FALSE, correct = TRUE)
  z <- qnorm(out$p.value / 2)
  r_eff <- abs(z) / sqrt(n)

  wilcox_results$Item[i]        <- sub(".*Q", "Q", js_col)
  wilcox_results$Section[i]     <- ifelse(grepl("JS1", js_col), "Section1",
     "Section2")
  wilcox_results$Median_JS[i]   <- median(js, na.rm = TRUE)
  wilcox_results$Median_ES[i]   <- median(es, na.rm = TRUE)
  wilcox_results$W_statistic[i] <- out$statistic
  wilcox_results$Z_statistic[i] <- z
  wilcox_results$p_value_raw[i] <- out$p.value
```

```
  wilcox_results$Effect_r[i]    <- r_eff
}

# Holm correction
wilcox_results$p_value_holm <- p.adjust(wilcox_results$p_value_raw, method =
    "holm")
wilcox_results$Sig <- ifelse(wilcox_results$p_value_holm < 0.05, "Yes", "No")

print(wilcox_results, row.names = FALSE)
write.csv(wilcox_results, "wilcox_holm_results_JS-ES.csv", row.names = FALSE)
```

Appendix C

Wilcoxon Signed-Rank Test Results With Holm Correction

Item	L1 *Mdn*	L2 *Mdn*	*w*	*z*	*p*	*p* (*Holm*)	*r*
S1Q1	2.00	4.00	96.00	−4.60	< .001	< .001	.65
S1Q2	2.00	3.00	7.50	−5.31	< .001	< .001	.75
S1Q3	2.00	3.00	7.00	−5.33	< .001	< .001	.75
S1Q4	2.00	4.00	25.50	−4.90	< .001	< .001	.69
S1Q5	3.00	4.00	46.00	−4.94	< .001	< .001	.70
S1Q6	2.00	3.00	54.00	−4.23	< .001	< .001	.60
S1Q7	2.00	4.00	13.00	−5.17	< .001	< .001	.73
S1Q8	2.00	4.00	7.00	−5.25	< .001	< .001	.74
S1Q9	2.00	4.00	22.50	−4.84	< .001	< .001	.68
S1Q10	2.00	3.00	55.00	−4.41	< .001	< .001	.62
S2Q1	2.00	4.00	37.50	−4.80	< .001	< .001	.68
S2Q2	3.00	4.00	100.00	−3.42	< .001	.002	.48
S2Q3	3.00	4.00	108.00	−3.79	< .001	< .001	.54
S2Q4	3.00	4.00	96.00	−4.08	< .001	< .001	.58
S2Q5	2.00	3.00	74.00	−3.61	< .001	.002	.51
S2Q6	2.00	3.00	105.00	−2.88	.004	.004	.41
S2Q7	2.00	4.00	75.50	−4.22	< .001	< .001	.60
S2Q8	3.00	4.00	122.50	−3.41	< .001	.002	.48
S2Q9	3.00	4.00	54.00	−4.32	< .001	< .001	.61
S2Q10	2.50	4.00	123.00	−3.50	< .001	.002	.49

Note. $N = 50$. S= Section; Q = Question; *Mdn* = Median; *p* (Holm) = Holm-adjusted *p*-values. L1 means Japanese, whereas L2 means English. Effect size (r) is evaluated as large (≥ 0.50), medium (≥ 0.30), and small (≥ 0.10) following Loerts et al. (2020). See Appendix A for questionnaire items and Appendix B for R command.