

Research Article

Revalidating the Foreign Language Enjoyment Scale in a new Context: A Confirmatory Factor Analysis Approach

Yohei Nakanishi^{id}^a, and Osamu Takeuchi^{id}^a^aKansai University, Japan

(Received 8 October 2025; accepted 8 July 2026)

Corresponding author: Yohei Nakanishi ✉

ABSTRACT

The Foreign Language Enjoyment Scale (FLES) has been developed and validated in various contexts; however, the applicability of a validated FLES from one context to another is not well established. To address this gap, the present study systematically examined the applicability of the FLES by evaluating models proposed in prior studies and investigating the validity and reliability of the adapted FLES, specifically in the Japanese English as a foreign language context. We employed theory-driven confirmatory factor analysis, analyzing data from 390 Japanese university students across six universities. Confirmatory factor analysis was applied to previously proposed two-factor and three-factor models of foreign language enjoyment that had been validated and used in different contexts. Our results indicated that three-factor models consistently exhibited better psychometric properties compared to the two-factor models. Among the models assessed, the model developed by Botes et al. (2021) demonstrated the best fit, showing that the previously validated scale was applicable to the Japanese English as a foreign language context. Modifications were also made to develop a refined model, and the results are discussed.

Keywords: foreign language enjoyment, Foreign Language Enjoyment Scale, positive psychology

INTRODUCTION

In recent years, the emphasis on the role of emotions in language learning in the field of applied linguistics has been on the rise (Dewaele, 2021). In the early stages of research in this field, most studies focused on negative emotions, particularly foreign language anxiety (FLA), as highlighted by Teimouri et al. (2019). Such initial studies were crucial in enhancing our understanding of how emotions affect language learning (Papi, 2010; Yan & Horwitz, 2008); however, the focus of emotion-related research changed significantly after the introduction of positive psychology into our field (MacIntyre et al., 2016). MacIntyre and Gregersen (2012) emphasized the importance of positive emotions in language learning and paved the way for new research into emotions. In response to this shift, Dewaele and MacIntyre (2014) introduced the concept of foreign language enjoyment (FLE) as a critical construct that embodies the notion of positive psychology. Since its inception, FLE has garnered significant scholarly attention, leading to a surge of studies examining positive emotions in language learning (Dewaele, Botes, et al., 2023; Elahi Shirvan et al., 2024; Li et al., 2018; Saito et al., 2018).

The theoretical foundation of FLE research is the broaden-and-build theory proposed by Fredrickson (2004). It theorizes that “positive emotions appear to broaden people’s momentary thought–action repertoires and build their enduring personal resources” (Fredrickson, 2004, p. 1369). Based on this theory, several studies have confirmed that high levels of FLE have been associated with increased motivation (Dewaele, Saito, et al., 2023), high engagement (Hosseini et al., 2022), improved second or foreign language (L2) speech development (Saito et al., 2018), increased willingness to communicate (Khajavy et al., 2018), and higher academic achievement (Jin & Zhang, 2021). As these studies show, academic interest in individual factors related to FLE is on the increase. To further advance FLE-related research, it is crucial to use validated and reliable research tools (Al-Hoorie et al., 2024). In response to this need, the Foreign Language Enjoyment Scale (FLES) has been adapted and validated

across a range of contexts, including different countries (e.g., Li et al., 2018) and educational levels (e.g., Shimizu et al., 2026).

Dewaele and MacIntyre (2014) developed the original FLES, a 21-item instrument designed to measure learners' enjoyment of foreign language learning. Given that emotions in language learning are culturally embedded (MacIntyre & Vincze, 2017), instruments such as the FLES need to be rigorously tested for validity and reliability across different cultural backgrounds. Li et al. (2018), for example, performed exploratory and confirmatory factor analyses to develop and validate the Chinese version of the FLES and assess its psychometric properties. Similarly, Nakanishi and Takeuchi (2024) validated the FLES for young language learners in the Japanese foreign language (FL) context. In addition, a short version of the FLES has also been rigorously validated (Botes et al., 2021). Previous studies have thus validated the FLES across different contexts and learner populations, building a foundation for FLE research. However, to our knowledge, limited attention has been paid to whether these validated models are applicable to new contexts.

Accordingly, this study aims to investigate whether previously validated FLE factor structures identified across different contexts and learner populations (e.g., Botes et al., 2021; Dewaele & MacIntyre, 2016; Li et al., 2018; Nakanishi & Takeuchi, 2024; Xethakis et al., 2025) are applicable to Japanese university students in the FL context. To this end, theory-driven confirmatory factor analysis (CFA) was employed to examine how well existing FLE models (e.g., Botes et al., 2021; Li et al., 2018; Nakanishi & Takeuchi, 2024) fit the present data based on model fit indices. To complement the CFA results and gain a deeper understanding of learners' FLE, qualitative data were also collected as part of a mixed-methods approach. By rigorously examining the applicability of the FLES across cultural contexts and learner populations, this study contributes empirical evidence to ongoing discussions regarding whether a context-dependent construct such as FLE can be validly measured in different settings.

LITERATURE REVIEW

Foreign Language Enjoyment

The study of FLE has received scholarly attention because of the increasing momentum of research in positive psychology (Botes et al., 2022a). Dewaele and MacIntyre (2016) defined FLE as “a complex emotion, capturing interacting dimensions of challenge and perceived ability that reflect the human drive for success in the face of difficult tasks” (p. 216). Furthermore, it is essential to distinguish enjoyment from pleasure (Csikszentmihalyi, 2008). Enjoyment occurs when individuals meet their basic needs and achieve new, unexpected, and interesting things, whereas pleasure is a simpler and more immediate feeling associated with enjoyable and likable experiences (Dewaele & MacIntyre, 2016).

The importance of enjoyment is well recognized in the field of positive psychology. According to flow theory, enjoyment is necessary to achieve a flow state, in which individuals become completely absorbed and immersed in certain activities (Csikszentmihalyi, 2008). Furthermore, the PERMA model (Positive emotions, Engagement, Relationships, Meaning, and Achievement) proposed by Seligman (2011) suggests that positive emotions, such as enjoyment, enhance people’s well-being. Similarly, the broaden-and-build theory proposed by Fredrickson (2004) speculates that “by broadening an individual’s momentary thought–action repertoire—whether through play, exploration or similar activities—positive emotions promote discovery of novel and creative actions, ideas and social bonds, which in turn build that individual’s personal resources” (Fredrickson, 2004, p. 1367).

Initially, FLE studies were conducted together with studies on FLA (Dewaele & MacIntyre, 2014). A recent meta-analysis by Botes et al. (2022a) indicated a moderate negative correlation ($r = -.31$) between FLA and FLE. However, it is important to note that the absence of FLA does not necessarily indicate the presence of FLE (Dewaele & MacIntyre, 2016). Numerous studies have investigated the factors that influence FLE and its reciprocal relationship with other variables. For instance, Dewaele, Saito, et al. (2023) examined the interaction between FLE, FLA, attitude, and motivation.

Their findings indicated that FLE and FLA were related to attitude and motivation. Similarly, Dewaele, Botes, et al. (2023) conducted a study on the relationship among FLE, foreign language classroom anxiety, and foreign language boredom and discovered that these variables were interrelated and influenced by learner factors, teacher behavior, and unpredictability. Lee (2022) examined the association between classroom enjoyment, grit, and willingness to communicate (WTC) and indicated that the perseverance of effort in grit and classroom enjoyment were predictors of WTC. Furthermore, Pawlak et al. (2024) investigated the interplay between academic emotions (enjoyment, boredom, and anxiety), self-perceived competence, L2 grit, and motivated behavior. The results of path analysis showed that enjoyment significantly predicted motivated behavior and L2 grit.

As these studies show, the growing body of research has integrated FLE as a key factor in language learning. This emerging trend highlights the need to develop and validate the FLES to accurately measure and understand the impact of FLE on L2 learning.

Measurement of Foreign Language Enjoyment

Pioneering researchers in the field of FLE, Dewaele and MacIntyre (2014), developed a scale based on Ryan et al.'s (1990) study. The 21-item FLES captures elements such as achievement, interest, and relationships with teachers and peers (Dewaele & MacIntyre, 2014). Numerous subsequent studies have attempted to validate the FLES based on this original scale. Table 1 summarizes these studies, including their references, participants, number of factors, factor names, and corresponding items for each factor.

Table 1. *Eight Versions of the Factor Structure and Corresponding Items of the Foreign Language Enjoyment Scale*

References	Participants	Number of factors	Factor name	Items ^a of FLES
Dewaele & MacIntyre (2016)	International	2	FLE-Social	10, 14, 15, 16, 17, 18, 19, 20, 21
			FLE-Private	3, 4, 8, 9, 12
Li et al. (2018)	Chinese	3	FLE-Private	3, 4, 8, 9, 12
			FLE-Teacher	15, 16, 17
			FLE-Atmosphere	10, 18, 19
Jin & Zhang (2021)	Chinese	3	Enjoyment of Teacher support	15, 16, 17
			Enjoyment of English learning	1, 3, 4, 5, 6, 7, 8, 9, 12
			Enjoyment of Student support	14, 18, 19, 20, 21
Ahmadi-Azad et al. (2020)	Iranian	3	FLE-Private	1, 4, 5, 6, 7, 8, 9
			FLE-Teacher	3, 15, 16, 17
			FLE-Atmosphere	10, 14, 18, 19
Botes et al. (2021)	International	3	Teacher Appreciation	15, 16, 17
			Personal Enjoyment	4, 8, 9
			Social Enjoyment	19, 20, 21
Saito et al. (2018)	Japanese	2	Social Enjoyment	10, 12, 14, 18, 21
			Private Enjoyment	3, 4, 7, 9, 11
Xethakis (2024)	Japanese	2	FLE-Social	14, 15, 16, 18, 19, 21
			FLE-Private	1, 4, 6, 7, 8, 9
Nakanishi & Takeuchi (2024)	Japanese Elementary	3	FLE-Private	1, 3, 6, 7, 8, 9, 13
			FLE-Teacher	15, 16, 17
			FLE-Atmosphere	10, 18, 19, 20, 21

Note. FLES = foreign language enjoyment scale; FLE = foreign language enjoyment.

^a The item numbers corresponding to each factor are based on Dewaele and MacIntyre (2016) and the original source studies, with additional reference to Xethakis (2024). For the full item descriptions, please refer to Appendix A.

Dewaele and MacIntyre (2016) conducted a study that involved 1,742 multilingual participants and employed principal component analysis with Varimax rotation for factor analysis. The factor analysis identified two distinct factors within the FLES. The first factor, FLE-Social, comprises nine items and represents enjoyment related to the social environment of language learning. The second factor, FLE-Private, consists of five items focused on individuals' personal enjoyment and intrinsic motivation in language learning. In subsequent studies, some researchers used the 21-item version developed by Dewaele and MacIntyre (2014), whereas others used the more economical 14-item version developed by Dewaele and MacIntyre (2016).

Li et al. (2018) created a Chinese version of the FLES, originally developed by Dewaele and MacIntyre (2016). This study was conducted with Chinese high school participants. The exploratory factor analysis (EFA) extracted three factors: FLE-Teacher, FLE-Private, and FLE-Atmosphere. In addition, the CFA result indicated that

the three-factor model with 11 items had a good fit, demonstrating $\chi^2(41) = 72.975$, comparative fit index (CFI) = .975, Tucker–Lewis index (TLI) = .967, standardized root mean square residual (SRMR) = .034, and root mean square error of approximation (RMSEA) = .041. These indices capture different aspects of model fit. Specifically, the CFI and TLI indicate how well the proposed model fits the data, with values closer to one generally suggesting a better fit, whereas the SRMR and RMSEA reflect the degree of model misfit, with lower values generally indicating a better fit (Hair et al., 2019). In another study conducted in China, Jin and Zhang (2021) investigated the dimensions of FLE using the original 21-item version developed by Dewaele and MacIntyre (2014). Using principal components analysis, this study identified three factors with 17 items: enjoyment of student support, enjoyment of teacher support, and enjoyment of English learning. In a subsequent study reported by Jin and Zhang (2019), the researchers conducted CFA, using a different dataset to validate the FLES developed by Jin and Zhang (2021), which had good psychometric properties with 16 items: $\chi^2(101) = 143.50$; CFI = .99; TLI = .99; GFI = .97; RMSEA = .03; SRMR = .05.

Ahmadi-Azad et al. (2020) used the 21-item FLES, originally designed and developed by Dewaele and MacIntyre (2014), in the Iranian FL context to explore how teachers' personalities influenced the FLE of 1,209 EFL learners. However, the reliability of the FLES was questioned based on the results of the pilot study. To improve the scale, CFA, EFA, and exploratory structural equation modeling were conducted, which resulted in the removal of six items with low factor loadings below 0.40. Consequently, a more robust 15-item scale with acceptable reliability scores was developed.

Botes et al. (2021) aimed to create and validate a short-form version of the FLES, as originally proposed by Dewaele and MacIntyre (2014). To achieve this, they used the dataset from the study conducted by Dewaele and MacIntyre (2014) and a rigorous methodology that included EFA with principal components analysis and Promax rotation, as well as algorithmic analysis, followed by CFA. The EFA identified three factors: Teacher Appreciation, Personal Enjoyment, and Social Enjoyment. Further,

CFA confirmed these findings with satisfactory fit indices: $\chi^2(24) = 89.25$, CFI = .978, TLI = .967, and RMSEA = .059.

In the Japanese FL context, Saito et al. (2018) used 10 items from the FLES by Dewaele and MacIntyre (2014) in their research design with a sample of 108 high school students. The researchers performed a factor analysis with Promax rotation and identified a two-factor model: Social Enjoyment and Private Enjoyment. Xethakis (2024) conducted a study with 514 Japanese university students to validate the Japanese version of the FLES based on the scale by Dewaele and MacIntyre (2014). This study initially employed CFA with maximum likelihood estimation. The data collected in this study were tested to determine whether they fit the four proposed FLE models (Botes et al., 2021; Jin & Zhang, 2021; Li et al., 2018; Saito et al., 2018). The analysis revealed that the model by Botes et al. (2021) was the best of all models; however, the fit indices were not satisfactory: $\chi^2(24) = 143.518$; CFI = .951; TLI = .926; SRMR = .057; RMSEA = .099. The researcher then conducted EFA and revealed a two-factor model, FLE-Social and FLE-Private. Nakanishi and Takeuchi (2024) also adapted and validated the 21-item scale developed by Dewaele and MacIntyre (2014) for young language learners using exploratory structural equation modeling, resulting in a 15-item version. Furthermore, additional studies have validated the 9-item short scale developed by Botes et al. (2021) in the Japanese context. Specifically, Xethakis et al. (2025) validated this short scale with university students, whereas Shimizu et al. (2026) validated it with young language learners.

As discussed above, FLE has been extensively examined, and the validity and reliability of the FLES have been tested across different cultures and learner populations. Some studies have proposed two-factor models of FLE (e.g., Dewaele & MacIntyre, 2016), and others three-factor models (e.g., Botes et al., 2021). Even among studies identifying the same number of factors, each factor comprised slightly different sets of items (e.g., Ahmadi-Azad et al., 2020; Jin & Zhang, 2021; Li et al., 2018). Given this growing body of research, it is now timely to examine whether an FLES validated in one cultural and educational setting can maintain its psychometric

properties, including validity and reliability, when applied in a different context. To address this issue, the present study aimed to test whether the existing FLE models discussed above are applicable to a Japanese EFL context. This study not only tests the applicability of the FLES in a new cultural setting but also enhances our understanding of how FLE may vary or remain consistent across diverse foreign language learning contexts.

The Current Study

The aim of this study was to examine the applicability of previously proposed two-factor and three-factor models of the FLES to the Japanese FL context by using CFA. Specifically, the study compared several existing versions of these models, which differed slightly in their item composition, to identify the structure that provides the most appropriate representation of FLE in the Japanese FL context. Furthermore, this study aimed to develop an empirically validated Japanese version of the FLES. To achieve this, the study addressed the following research question:

RQ: To what extent do the previously proposed two-factor and three-factor models of FLE, including their different item compositions, demonstrate evidence of validity and reliability in the Japanese FL context?

RESEARCH METHODS

Participant Recruitment and Sampling

The convenience sampling method was employed; however, in order to enhance the applicability to Japanese university students learning English as a foreign language, we made sure to include participants with a wide range of academic backgrounds and from different regions across Japan in our data collection based on Andringa and Godfroid's (2020) recommendation. Furthermore, as teacher factors seem to influence FLE (Botes et al., 2021; Li et al., 2018), the present study made efforts to include as many English teachers of the study participants as possible to minimize the influence

of any single teacher. As a result, we collected data from students at six different universities located in western, central, and eastern parts of Japan, and seven teachers collaborated to collect the data during their classes. The participants were from six universities: University A ($n = 54$), University B ($n = 27$), University C ($n = 119$), University D ($n = 75$), University E ($n = 40$), and University F ($n = 73$). Two participants did not provide their university name. The participants' academic backgrounds were diverse. They were majoring in economics, social safety sciences, law, international studies, and other disciplines. Of the 390 participants, 211 were men, 174 were women, and five did not disclose their gender. Participants were generally between 18 and 22 years old, the typical age range for university students in Japan.

Instruments

Foreign Language Enjoyment Scale

A Japanese version of the FLES was adapted from the original FLES developed by Dewaele and MacIntyre (2014). The original FLES was translated into Japanese by the first author for the present study, with some phrasing also informed by previous studies (Saito et al., 2018; Xethakis, 2024). Subsequently, three researchers in our field reviewed the translated version for word choice, comprehensibility, and consistency with the original version. Further, the scale was translated back into English by an English and Japanese bilingual researcher in the field of applied linguistics. Disagreements were resolved through discussion. The following is an example of how the English items were adapted for the Japanese context: Item 15, “The teacher is encouraging” was translated as “英語の先生はいつも励まそうとしてくれる.” Key linguistic adaptations included specifying “英語の先生” (English teacher) for contextual clarity, translating “encouraging” as “励まそうとしてくれる” to capture the intentional effort implied in the original item, and adding “いつも” (always) to express the continuous nature of the English present tense. A pilot test was conducted with 30 university students, and

some wording was adjusted based on the feedback. The pilot test showed that the scale had a high level of reliability with a Cronbach's alpha coefficient of .913.

Open-Ended Question

This study included the following open-ended question to triangulate the findings from the quantitative data and to ensure consistency with the study conducted by Dewaele and MacIntyre (2016): "Please describe a specific event (activities) or episode in your EFL class that you really enjoyed, and describe your feelings in as much detail as possible." The question was also translated into Japanese to enable participants to express their feelings accurately.

Data Collection and Ethical Considerations

The data were collected in October and November of 2023. During this period, participants completed the Japanese version of the FLES in their English classes via Google Forms. Prior to data collection, we received approval for this research from our institution's ethics review committee. Before conducting the study, the participants were informed in a written form that their data would be kept anonymous and secure. The collected data would not be used to grade or evaluate them, and only researchers would have access to the data. After this explanation, the questionnaire began with a consent form for participants. Only data from individuals who agreed to participate were used in the subsequent analysis.

Data Analyses

All analyses were conducted using JASP (Version 0.18; JASP Team, 2023). To address the research question, we performed CFA with maximum likelihood estimation on eight previously proposed scale versions discussed in the literature review. These versions included both two-factor and three-factor structures, with differences mainly in item composition across studies. ML estimation, rather than maximum likelihood robust estimation, was selected because the sample size was

sufficiently large relative to the number of measured variables, there were no missing data, and the data did not show substantial non-normality (Hair et al., 2019). Specifically, the selected scale versions included the two-factor model proposed in the pioneering study by Dewaele and MacIntyre (2016), versions validated in specific contexts (Ahmadi-Azad et al., 2020; Jin & Zhang, 2021; Li et al., 2018), a short-scale version (Botes et al., 2021), and versions previously used in the Japanese context (Nakanishi & Takeuchi, 2024; Saito et al., 2018; Xethakis, 2024). Although EFA is commonly used in questionnaire development, CFA was chosen for this study given the existence of robust, theoretically grounded models of FLE (Plonsky & Gonulal, 2015). This approach allowed us to directly examine whether these established models (e.g., Botes et al., 2021; Jin & Zhang, 2021; Li et al., 2018) were applicable in the present context.

To assess how well the data fit the models, model fit was determined using various criteria (Hair et al., 2019; In'nami & Koizumi, 2011): χ^2 , RMSEA ($< .07$), SRMR ($< .08$), GFI ($> .95$), CFI ($> .95$), TLI ($> .95$). Furthermore, other fit indices such as Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were assessed. These indices were used to evaluate which of the eight previously proposed scale versions best fit the data collected in the Japanese FL context. Generally, models with lower AIC and BIC scores are preferred (Kline, 2023). Additionally, we conducted a reliability test to assess the internal consistency using both Cronbach's alpha and McDonald's omega. While Cronbach's alpha is widely used, McDonald's omega is known to be more robust and reliable (Hayes & Coutts, 2020). Convergent validity was analyzed by examining the average variance extracted (AVE), which was required to exceed 0.50. Additionally, discriminant validity was evaluated using the Heterotrait-Monotrait ratio (HTMT) with a cutoff criterion set at 0.85 (Hair et al., 2019).

To support the quantitative findings, we conducted a qualitative analysis to examine whether the scale results aligned with participants' experiences and feelings described in their open-ended responses. The enjoyable episodes reported by participants were categorized by the first author according to the three-factor model of FLE (Teacher

Appreciation, Social Enjoyment, and Personal Enjoyment) established by Botes et al. (2021), which demonstrated the best fit in the Japanese EFL context through CFA. Then, the first author re-checked all responses after an interval to ensure analytical accuracy. For transparency and cautious judgment, this qualitative analysis was conducted after the CFA identified the three-factor model as the best-fitting framework, which may present a risk of confirmation bias in the coding process. To address this potential concern, the coding focused primarily on the semantic content of participants' responses rather than forcing them to fit the predetermined categories. For instance, a participant's statement, such as "I find that I can improve my comprehension by exchanging opinions and answers with my friends during group work," indicates exchanging opinions, friends, and group work. These terms reflect the social element of FLE; therefore, this statement was categorized as social enjoyment.

RESULTS

Descriptive Statistics

In this study, 395 responses without missing data were collected using Google Forms. However, five responses were excluded because informed consent was not obtained from those participants at the beginning of the questionnaire, as described in the Data Collection and Ethical Considerations section. Finally, 390 valid records were obtained for analysis. The descriptive statistics showed that the average item mean and average standard error were 3.583 and 0.051, respectively. In addition, the average skewness and kurtosis values, which were examined to assess the data distribution, were -0.565 and 0.395, respectively.

Confirmatory Factor Analysis

This section is structured in three stages. First, we present the results of the eight model versions to identify the best-fitting factor structure. Second, we examine the

psychometric properties of the best-fitting model. Third, we present the results of a refined model in the Japanese EFL context.

This study used CFA with ML estimation to assess the fit of the collected data to the previously proposed scale versions (see Table 1). Table 2 summarizes the results. Initially, the data were applied to the two-factor model developed by Dewaele and MacIntyre (2016); however, the result showed a poor fit. Subsequently, the data were tested against two scale versions developed in the Chinese FL context by Li et al. (2018) and Jin and Zhang (2021). The fit indices of the three-factor model by Li et al. (2018) were barely acceptable except for RMSEA. The result of the three-factor model by Jin and Zhang (2021) was a poor fit. Thereafter, CFA was performed on the three-factor model used in the Iranian FL context by Ahmadi-Azad et al. (2020). The result was also a poor fit. Further analyses were conducted on the three-factor model proposed by Botes et al. (2021). It exhibited satisfactory fit indices. Finally, CFA was conducted on three scale versions developed in the Japanese context by Saito et al. (2018), Xethakis (2024) and Nakanishi and Takeuchi (2024). The analysis of all models yielded poor fit indices. Overall, the three-factor models (Ahmadi-Azad et al., 2020; Botes et al., 2021; Jin & Zhang, 2021; Li et al., 2018) consistently demonstrated a better fit than the two-factor models (Dewaele & MacIntyre, 2016; Saito et al., 2018; Xethakis, 2024). We compared the AIC and BIC values to identify the best-fitting model for the data gathered in the Japanese FL context. The model proposed by Jin and Zhang (2021) scored 15507 on AIC and 15721 on BIC, which were the highest scores, whereas the model by Botes et al. (2021) had the lowest scores of 8095 and 8214 on AIC and BIC, respectively. Because lower scores are better in these comparative indices, we concluded that the model proposed by Botes et al. (2021) was the best fit for the data.

Table 2. *Summary of Fit Indices for Each Model*

References	Factor model	χ^2	CFI	TLI	GFI	RMSEA		SRMR	AIC	BIC
						value	90% CI			
Dewaele & MacIntyre (2016)	2	569.04*	.856	.827	.959	.129	[.119, .139]	.083	12435	12606
Li et al. (2018)	3	123.00*	.970	.960	.993	.072	[.057, .086]	.046	9189	9332
Jin & Zhang (2021)	3	434.68*	.916	.901	.973	.084	[.076, .092]	.060	15507	15721
Ahmadi-Azad et al. (2020)	3	269.63*	.943	.931	.985	.073	[.064, .083]	.078	13399	13589
Botes et al. (2021) ^a	3	67.29*	.977	.965	.996	.068	[.049, .088]	.044	8095	8214
Saito et al. (2018)	2	298.73*	.840	.789	.981	.141	[.127, .156]	.076	9750	9873
Xethakis (2024)	2	267.79*	.910	.888	.982	.102	[.090, .114]	.058	11046	11193
Nakanishi & Takeuchi (2024)	3	388.77*	.905	.885	.866	.094	[.088, .104]	.083	13486	13617

Note. χ^2 = chi-squared statistic; CFI = comparative fit index; TLI = Tucker-Lewis index; GFI = goodness-of-fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion. * $p < .001$.

^aThis model showed the best fit.

We identified the model developed by Botes et al. (2021) as the best-fitting model; therefore, a detailed analysis was conducted. Table 3 presents the comprehensive results of factor loadings and residual variances. Regarding factor loadings, each item for the intended factor showed high values ranging from 0.715 to 0.861, demonstrating a well-defined factor structure of FLE. Regarding residual variance, Item 20 had a value of 0.99, indicating that this item failed to capture social enjoyment in the Japanese foreign language context. We acknowledge that our decision to remove Item 20 represented a post-hoc, data-driven modification that deviates from the theoretical foundations of CFA, which should ideally be guided by theoretical considerations rather than statistical outcomes alone. However, the exceptionally poor residual

variance of Item 20 was further supported by our following qualitative analysis of open-ended questionnaire responses, in which no participants mentioned the content of Item 20 (“We have common ‘legends,’ such as running jokes”). In addition, the English term “legends” in this context was challenging to translate accurately into Japanese, as it refers to informal, shared stories or inside jokes that might not have a direct cultural equivalent in Japanese educational contexts. Combining the extremely high residual variance with the absence of references to this item’s content in our qualitative data and the translation challenges, we determined that Item 20 was not appropriate for the Japanese foreign language context. Therefore, Item 20 was removed, and the model fit indices were recalculated for the refined model adapted to the Japanese context. Table 4 summarizes the model fit indices of both the refined model adapted to the Japanese context and Botes et al.’s (2021) original model, demonstrating that all fit indices except GFI were further improved in the refined model. Table 5 presents the detailed CFA results for the refined FLES model. Figure 1 illustrates the model plot of the refined model adapted to the Japanese context, demonstrating a well-structured model with good factor loadings and acceptable residual variances.

Table 3. *Detailed Information of the Model by Botes et al. (2021), Including Item 20*

Factor	Item	Factor loading	Residual variance
Teacher Appreciation	15	0.757	0.225
	16	0.756	0.079
	17	0.754	0.116
Personal Enjoyment	4	0.804	0.490
	8	0.747	0.510
	9	0.715	0.475
Social Enjoyment	19	0.861	0.307
	20	0.736	0.990
	21	0.838	0.482

Table 4. Model Fit Indices for Refined Model of Botes et al. (2021), Excluding Item 20

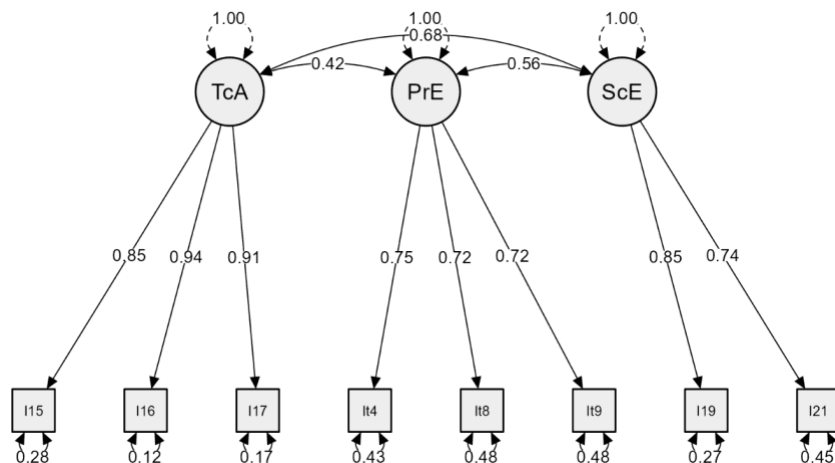
Model	χ^2	CFI	TLI	GFI	RMSEA		SRMR	AIC	BIC
					value	90% CI			
Refined model based on Botes et al. (2021)	30.69	.992	.987	.981	.045	[.017, .071]	.036	6932	7008
Botes et al. (2021)	67.29	.977	.965	.996	.068	[.049, .088]	.044	8095	8214

Note. χ^2 = chi-squared statistic; CFI = comparative fit index; TLI = Tucker-Lewis index; GFI = goodness-of-fit index; RMSEA = root mean square error of approximation; CI = confidence interval; SRMR = standardized root mean square residual; AIC = Akaike information criterion; BIC = Bayesian information criterion.

Table 5. Details of Confirmatory Factor Analysis for the Refined Model

Factor	Item	Std. estimate	Std. Error	z-value	<i>p</i>	90% CI [Lower, Upper]	
Teacher	15	0.848	0.017	51.347	< .001	0.815	0.880
Appreciation	16	0.937	0.011	88.315	< .001	0.916	0.958
	17	0.912	0.012	75.635	< .001	0.888	0.935
Personal	4	0.753	0.033	22.716	< .001	0.688	0.819
Enjoyment	8	0.722	0.034	21.001	< .001	0.654	0.789
	9	0.722	0.034	21.042	< .001	0.655	0.790
Social	19	0.854	0.029	29.097	< .001	0.796	0.911
Enjoyment	21	0.740	0.032	23.043	< .001	0.677	0.803

Note. CI = confidence interval.

Figure 1. Refined Three-Factor Model of Foreign Language Enjoyment: Teacher Appreciation, Personal Enjoyment, and Social Enjoyment

Note. TcA = Teacher Appreciation, PrE = Personal Enjoyment, and ScE = Social Enjoyment.

Construct, Convergent, and Discriminant Validity and Reliability

Construct validity was confirmed based on the model fit indices (see Table 4) and the factor loadings for intended factors (see Table 5 and Figure 1). To examine convergent validity, AVE was used. The AVE of each factor model (Table 6) was greater than .50, which indicates good convergent validity (Kline, 2023). Discriminant validity was assessed using HTMT. Each value summarized in Table 7 was lower than .85, confirming good discriminant validity (Hair et al., 2019). Table 8 summarizes the reliability of each model. The refined model had McDonald's ω of .904 and Cronbach's α of .853, demonstrating good reliability.

Table 6. Average Variance Extracted for Each Factor

Factor	AVE
Teacher Appreciation	0.803
Personal Enjoyment	0.538
Social Enjoyment	0.633

Note. AVE = average variance extracted.

Table 7. Results of Heterotrait-Monotrait Ratio (HTMT)

Teacher Appreciation	Personal Enjoyment	Social Enjoyment
1		
0.414	1	
0.692	0.564	1

Table 8. Reliability of the Refined Model

Factor	Coefficient ω	Coefficient α
Teacher Appreciation	0.926	0.924
Personal Enjoyment	0.777	0.776
Social Enjoyment	0.775	0.773
Total	0.904	0.853

Qualitative Analysis of Open-Ended Responses

We analyzed qualitative data to gain a deeper understanding of how participants experienced FLE in the language classroom and to examine whether their open-ended responses were broadly consistent with the quantitative findings. Among 390 participants, 262 responded to the open-ended question. Some participants provided incomplete answers to the open-ended question, such as simply stating “group work.” Consequently, these responses were excluded from the analysis. In total, there were seven unclassified responses. Table 9 summarizes the sources of enjoyment mentioned by the participants more than 10 times. For this study, we assigned a unique identifier, such as A001 and A390 to each participant to ensure their privacy when we used their episodes as a direct citation.

Table 9. *Sources of Enjoyment Mentioned by Students*

	Personal enjoyment	Social enjoyment	Teacher appreciation
Sources	realization of progress, expressing oneself, a sense of accomplishment, authentic use of foreign language, interest and novelty, having fun, and others	learning from others, peer communication, building bonds, a good atmosphere, and others	pedagogical practice, and others
Frequency	124	117	14

Note. This table includes all sources of enjoyment mentioned more than 10 times. Responses mentioned fewer than 10 times were grouped as “Others” because they were highly individual and did not form clear, common categories.

Personal Enjoyment

Research indicates that personal enjoyment is a crucial factor in foreign language learning. The analysis suggested that six subcategories of personal enjoyment emerged from the data: realization of progress, expressing oneself, a sense of

accomplishment, authentic use of a foreign language, interest and novelty, and having fun.

The realization of progress was frequently mentioned. An example statement for this subcategory is as follows: “I feel happy about my English skills improving when I can express my thoughts in my own words” (A160). Additionally, the statement “Learning new idioms makes me feel like I am expanding my vocabulary and getting smarter” (A301) supports this theme.

Expressing oneself in a foreign language is a significant aspect of personal enjoyment in the context of English as a foreign language (EFL). An example statement reflecting this aspect is as follows: “When I talk about myself in English, it’s really enjoyable to have others understand and listen to my stories” (A077).

In EFL learning, students experience a sense of accomplishment; however, the factors leading to this varied in our data. Some students found a sense of achievement while participating in group activities with their peers, whereas others may have derived it from doing well on quizzes. This is demonstrated by the following statements: “When I participate in group discussions or presentations, my classmates praise me, and I feel a sense of accomplishment” (A033), and “When I successfully complete English quizzes, it gives me a sense of achievement” (A319).

In EFL classrooms, learners often have limited opportunities to use authentic language. Therefore, they may value opportunities to engage in English conversations to improve their language skills. This is exemplified by the following statement: “Despite my limited knowledge, when I speak English without using Japanese, it is a great feeling to have my peers understand what I am saying” (A360). This highlights learners’ appreciation for using authentic language despite facing linguistic challenges. Finally, learning new and interesting things is a source of enjoyment, as indicated by the following statement: “When encountering something new, it satisfies my curiosity and desire for knowledge” (A087).

Their statements reflected the items of the FLES, such as “I enjoy it,” “I’ve learned interesting things,” and “In class, I feel proud of my accomplishments,” indicating that the open-ended responses were broadly consistent with the scale content.

Social Enjoyment

The analysis revealed that social interactions with peers were an important source of enjoyment in learning environments and identified four key themes: learning from others, peer communication, building bonds, and a good atmosphere.

The most common theme that appeared was learning from others. Some representative perspectives were as follows: “I find that I can improve my comprehension by exchanging opinions and answers with my friends during group work” (A320), and “It’s enjoyable to hear different perspectives and collaborate with my peers during discussions” (A012).

The analysis indicated that peer communication is a significant source of social enjoyment. This aspect was often expressed as follows: “I enjoy having conversations with my friends and talking about different things in English” (A359). Furthermore, peer interactions during English lessons can improve human relationships and create new friendships. An example of this idea is as follows: “By sharing my opinions in a group, I have made many friends and have become close to them thanks to the English classes” (A279).

The key to social enjoyment seems to be establishing a good atmosphere. One example of such an environment is as follows: “It is very comfortable and relaxing when I think and present together” (A322), emphasizing the significance of realizing a positive and cooperative learning environment.

The items of the FLES such as “We form a tight group” and “We laugh a lot” accurately reflected participants’ experiences of social enjoyment in the Japanese FL context.

Teacher Appreciation

We noted that personal and social enjoyment were mentioned more frequently than teacher appreciation (see Table 9). However, some participants indicated that better teaching practices increased enjoyment. For instance, one participant said, “When I take part in quiz-like mini-games in groups, I don’t feel any pressure to speak formal English or strictly follow grammatical rules that are typically associated with studying” (A218). This comment highlights the significance of adopting relaxed and engaging teaching techniques to improve student enjoyment.

The role of teachers in establishing fun and enjoyable learning environments in foreign language classrooms is critical. One participant stated, “When the teacher is always smiling and creates a positive atmosphere in the classroom, it fosters an easy-to-learn and peaceful environment. Students feel comfortable asking questions” (A217). This statement highlights the significance of a supportive and positive classroom atmosphere maintained by the teacher. These comments from the participants provided further support for the contents of the FLES, such as “The teacher is encouraging,” “The teacher is friendly,” and “The teacher is supportive.”

DISCUSSION

To address the research question and determine whether a context-dependent emotion such as FLE can be accurately measured in the Japanese FL context, the present study tested previously proposed FLE models. These models reflected two broad factor structures, namely two-factor and three-factor models, and included eight versions with varying item compositions across studies. The theory-driven CFA consistently demonstrated that the three-factor model yielded better fit indices. This might suggest the benefit of separating the teacher’s role from the social component of FLE and treating it as a distinct factor. This is in line with the large body of literature in SLA emphasizing the teacher’s critical importance in the foreign language classroom (Ahmadi-Azad et al., 2020; Dewaele et al., 2019; Piechurska-Kuciel, 2011).

Among the three-factor models, the model developed by Botes et al. (2021) best fit our collected data. This finding is also consistent with subsequent studies that validated the short-scale version developed by Botes et al. (2021) with Japanese university students (Xethakis et al., 2025) and elementary school students (Shimizu et al., 2026). It is also consistent with an earlier study by Xethakis (2024), which also concluded that the model by Botes et al. (2021) was the best fit among the four models, although the fit indices of his study were not acceptable and thus the results cannot be used as definitive support for the superiority of the model by Botes et al. (2021). This may be because our research aimed to include participants from various universities and academic backgrounds. We believe that the diversity of participants contributed to this result, suggesting that, in addition to sample size, diversity played an equally important role in investigating FLE. This is because emotions such as FLE are highly context-dependent (MacIntyre & Vincze, 2017), and how foreign language learners experience language-related emotions varies across different academic backgrounds, instructional styles, and learning environments.

However, this study also revealed challenges in applying the established FLE model to different contexts. Specifically, Item 20, “We have common ‘legends,’ such as running jokes,” showed high residual variance (see Table 3), indicating that this item did not effectively measure the intended factor as mentioned above. Although the model fit indices improved, further research is necessary to explore the possibility of incorporating culturally grounded items that capture sources of enjoyment unique to the Japanese EFL context, especially social enjoyment, rather than only removing items that might not fit well with the cultural context.

In our study, we conducted a qualitative analysis of participants’ enjoyable episodes to examine whether their descriptions were broadly consistent with the FLES in the Japanese FL context. This analysis identified various sources of enjoyment that were largely consistent with those identified by Li et al. (2018) and Xethakis (2024), supporting the broad applicability of FLE across different contexts. However, our analysis also revealed some notable discrepancies between the frequency of enjoyment sources in the qualitative data and their representation in the final scale

structure. Sources of social enjoyment, such as peer communication, were frequently mentioned by participants but became less represented in the FLES following the removal of Item 20. Meanwhile, teacher appreciation appeared in only 14 of the 262 responses, despite being retained as a factor. One possible explanation is that the items in this subscale, such as Item 17 (“The teacher is supportive”), may reflect predictors of enjoyment, such as teacher support or encouragement, rather than enjoyment itself. From this perspective, the discrepancy may reflect not simply a gap between participants’ reported experiences and the scale structure, but also a conceptual mismatch between their descriptions of enjoyable classroom experiences and what this subscale measures. This possibility highlights the need for future research to examine more carefully how qualitative data, such as participants’ open-ended responses and interview data, can inform scale development and help determine whether some subscales represent enjoyment itself or its potential predictors.

CONCLUSION

Given the growing body of research, the present study employed theory-driven CFA to evaluate previously proposed models and examine the applicability of the FLES in the Japanese FL context. Our findings indicated that the three-factor model developed by Botes et al. (2021) provided the best fit to the Japanese FL context. This result is in line with findings from studies that validated the short-scale version with Japanese elementary school students (Shimizu et al., 2026) and university students (Xethakis et al., 2025). In the present study, however, Item 20 did not perform well. After removing Item 20, the refined model based on Botes et al. (2021) showed improved fit indices in the Japanese FL context. In other words, the shorter version outperformed the longer one. Furthermore, qualitative analysis of the open-ended responses generally reflected the content of the FLES items.

Our findings provide important insights for the applicability of FLE. This study demonstrated the possibility that scales developed in different contexts—specifically the one developed by Botes et al. (2021)—can be applied to the Japanese FL context. It should be noted that numerous factors influence the applicability of an original scale to different contexts, such as participant diversity, number of items, and methodological rigor in the original study.

Our study also has important theoretical and practical implications. From a theoretical perspective, this research demonstrates how validated scales can be applied in diverse contexts. The findings suggest that established measurement tools for other emotional constructs, such as foreign language boredom (Pawlak et al., 2020), may benefit from similar cross-cultural validation efforts. This study also provides guidance for researchers working with validated scales across various contexts. Even when using well-established instruments, researchers need to carefully assess both validity and reliability within their own cultural and linguistic contexts.

From a practical perspective, this Japanese version of the FLES can be used to evaluate FLE efficiently owing to its concise eight-item format. The short format of the Japanese FLES makes it relatively easy to manage, thereby lowering the burden on learners, researchers, and practitioners, while it allows them to evaluate students' FLE levels accurately. Based on these assessments, teachers can develop informed teaching plans and select appropriate lesson types. When designing teaching and lesson plans to foster their FLE in the Japanese EFL context, it may be helpful to gain some ideas from, for example, Fresacher (2016), who explained how to apply positive psychology-based activities in the classroom.

Before concluding, it is important to acknowledge some limitations of this study. First, our study had limited data on participants' backgrounds, including their English proficiency levels, previous learning experiences, and specific classroom environments. Nor did we consider the potential influence of teachers' instructional styles on students' FLE. The lack of such measures restricts the interpretation of our CFA results and the broader applicability of our findings. Therefore, further research

is needed to explore these relationships in line with recent studies by Botes et al. (2022b) and Yang and Lin (2026), using the valid and reliable scale developed in this study. Additionally, future studies should examine the predictive validity of the Japanese FLES by investigating correlations with key variables such as motivation and achievement. Second, the Japanese version of the FLES was examined using only data from university students, which limits its applicability to other groups of language learners, such as high school students in the Japanese EFL context. Therefore, it is necessary to gather data from high school and junior high school students to further examine the applicability of the Japanese version of the FLES. Additionally, this study applied CFA to previously proposed FLE scale versions in the Japanese context. To confirm the extent to which findings hold in the broader Japanese EFL context, further efforts are needed using a replication framework (McManus, 2024).

Despite the aforementioned limitations, this study addresses the context-specific nature of emotions such as FLE by examining the applicability of previously proposed versions of the FLES in the Japanese FL context. It highlights the importance of testing existing scales within a specific context before using them in further research. In this sense, the study contributes to a broader understanding of how FLE can be understood and measured across contexts.

Authors' Contributions

YN and OT participated in the design of the study and completed the data collection. YN worked on data analysis. YN and OT drafted the manuscript and participated in the interpretation of the results. Both authors read and approved the final manuscript.

Acknowledgments

We would like to thank all the teachers and participants involved in this research.

Declaration of GenAI and AI-Assisted Technologies

In the preparation of this work, the author(s) used Grammarly and Claude Opus 4.8 and earlier models in order to improve the clarity and consistency of language usage. After using these tools, the authors reviewed and edited the content as needed,

verified the accuracy, originality, and appropriateness of any GenAI-assisted outputs, and take full responsibility for the content of the publication.

Ethics Approval and Consent to Participate

This study was approved by the University Research Ethics Committee (approval no. 23-35). All participants provided written informed consent prior to enrollment and data collection in the study.

Funding

This study was supported by JST SPRING, Grant Number JPMJSP2150, granted to the first author of this article.

REFERENCES

- Ahmadi-Azad, S., Asadollahfam, H., & Zoghi, M. (2020). Effects of teacher's personality traits on EFL learners' foreign language enjoyment. *System*, 95, Article 102369. <https://doi.org/10.1016/j.system.2020.102369>
- Al-Hoorie, A. H., Hiver, P., & In'nami, Y. (2024). The validation crisis in the L2 motivational self system tradition. *Studies in Second Language Acquisition*, 46(2), 307–329. <https://doi.org/10.1017/S0272263123000487>
- Andringa, S., & Godfroid, A. (2020). Sampling bias and the problem of generalizability in applied linguistics. *Annual Review of Applied Linguistics*, 40, 134–142. <https://doi.org/10.1017/S0267190520000033>
- Botes, E., Dewaele, J.-M., & Greiff, S. (2021). The development and validation of the short form of the foreign language enjoyment scale. *The Modern Language Journal*, 105(4), 858–876. <https://doi.org/10.1111/modl.12741>
- Botes, E., Dewaele, J.-M., & Greiff, S. (2022a). Taking stock: A meta-analysis of the effects of foreign language enjoyment. *Studies in Second Language Learning and Teaching*, 12(2), 205–232. <https://doi.org/10.14746/ssllt.2022.12.2.3>
- Botes, E., Dewaele, J.-M., & Greiff, S. (2022b). By the old Gods and the new: The effect of the congruence and incongruence of foreign language classroom anxiety and enjoyment on self-perceived proficiency. *The Modern Language Journal*, 106(4), 784–797. <https://doi.org/10.1111/modl.12807>
- Csikszentmihalyi, M. (2008). *Flow: The psychology of optimal experience*. Harper Perennial.
- Dewaele, J.-M. (2021). The emotional rollercoaster ride of foreign language learners and teachers: Sources and interactions of classroom emotions. In M. Simons & T. F. H. Smits (Eds.), *Language education and emotions* (pp. 207–222). Routledge.
- Dewaele, J.-M., Botes, E., & Greiff, S. (2023). Sources and effects of foreign language enjoyment, anxiety, and boredom: A structural equation modeling approach. *Studies in*

Second Language Acquisition, 45(2), 461–479.

<https://doi.org/10.1017/S0272263122000328>

- Dewaele, J.-M., & MacIntyre, P. D. (2014). The two faces of Janus? Anxiety and enjoyment in the foreign language classroom. *Studies in Second Language Learning and Teaching*, 4(2), 237–274. <https://doi.org/10.14746/ssllt.2014.4.2.5>
- Dewaele, J.-M., & MacIntyre, P. D. (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). Multilingual Matters. <https://doi.org/10.21832/9781783095360-010>
- Dewaele, J.-M., Magdalena, A. F., & Saito, K. (2019). The effect of perception of teacher characteristics on Spanish EFL learners' anxiety and enjoyment. *The Modern Language Journal*, 103(2), 412–427. <https://doi.org/10.1111/modl.12555>
- Dewaele, J.-M., Saito, K., & Halimi, F. (2023). How foreign language enjoyment acts as a buoy for sagging motivation: A longitudinal investigation. *Applied Linguistics*, 44(1), 22–45. <https://doi.org/10.1093/applin/amac033>
- Elahi Shirvan, M., Taherian, T., & Yazdanmehr, E. (2024). Foreign language enjoyment: A longitudinal confirmatory factor analysis-curve of factors model. *Journal of Multilingual and Multicultural Development*, 45(2), 224–242. <https://doi.org/10.1080/01434632.2021.1874392>
- Fredrickson, B. L. (2004). The broaden-and-build theory of positive emotions. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, 359(1449), 1367–1377. <https://doi.org/10.1098/rstb.2004.1512>
- Fresacher, C. (2016). Why and how to use positive psychology activities in the second language classroom. In P. MacIntyre, T. Gregersen, & S. Mercer (Eds.), *Positive psychology in SLA* (pp. 344–358). Multilingual Matters. <https://doi.org/10.21832/9781783095360-016>
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2019). *Multivariate data analysis* (8th ed.). Cengage Learning.
- Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But.... *Communication Methods and Measures*, 14(1), 1–24. <https://doi.org/10.1080/19312458.2020.1718629>
- Hosseini, H. M., Fathi, J., Derakhshesh, A., & Mehraein, S. (2022). A model of classroom social climate, foreign language enjoyment, and student engagement among English as a foreign language learners. *Frontiers in Psychology*, 13, Article 933842. <https://doi.org/10.3389/fpsyg.2022.933842>
- In'nami, Y., & Koizumi, R. (2011). Structural equation modeling in language testing and learning research: A review. *Language Assessment Quarterly*, 8(3), 250–276. <https://doi.org/10.1080/15434303.2011.582203>
- JASP Team (2023). JASP (Version 0.18) [Computer software].

- Jin, Y., & Zhang, L. J. (2019). A comparative study of two scales for foreign language classroom enjoyment. *Perceptual and Motor Skills*, 126(5), 1024–1041. <https://doi.org/10.1177/0031512519864471>
- Jin, Y., & Zhang, L. J. (2021). The dimensions of foreign language classroom enjoyment and their effect on foreign language achievement. *International Journal of Bilingual Education and Bilingualism*, 24(7), 948–962. <https://doi.org/10.1080/13670050.2018.1526253>
- Khajavy, G. H., MacIntyre, P. D., & Barabadi, E. (2018). Role of the emotions and classroom environment in willingness to communicate: Applying doubly latent multilevel analysis in second language acquisition research. *Studies in Second Language Acquisition*, 40(3), 605–624. <https://doi.org/10.1017/S0272263117000304>
- Kline, R. B. (2023). *Principles and practice of structural equation modeling (5th ed.)*. Guilford Publications.
- Lee, J. S. (2022). The role of grit and classroom enjoyment in EFL learners' willingness to communicate. *Journal of Multilingual and Multicultural Development*, 43(5), 452–468. <https://doi.org/10.1080/01434632.2020.1746319>
- Li, C., Jiang, G., & Dewaele, J.-M. (2018). Understanding Chinese high school students' foreign language enjoyment: Validation of the Chinese version of the foreign language enjoyment scale. *System*, 76, 183–196. <https://doi.org/10.1016/j.system.2018.06.004>
- MacIntyre, P., & Gregersen, T. (2012). Emotions that facilitate language learning: The positive-broadening power of the imagination. *Studies in Second Language Learning and Teaching*, 2(2), 193–213. <https://doi.org/10.14746/ssllt.2012.2.2.4>
- MacIntyre, P. D., Gregersen, T., & Mercer, S. (Eds.). (2016). *Positive psychology in SLA. Multilingual Matters*. <https://doi.org/10.21832/9781783095360>
- MacIntyre, P. D., & Vincze, L. (2017). Positive and negative emotions underlie motivation for L2 learning. *Studies in Second Language Learning and Teaching*, 7(1), 61–88. <https://doi.org/10.14746/ssllt.2017.7.1.4>
- McManus, K. (2024). Replication studies in second language acquisition research: Definitions, issues, resources, and future directions: Introduction to the special issue. *Studies in Second Language Acquisition*, 46(5), 1299–1319. <https://doi.org/10.1017/s0272263124000652>
- Nakanishi, Y., & Takeuchi, O. (2024). Validating the foreign language enjoyment scale for young learners: An exploratory structural equation modeling approach. *Research Methods in Applied Linguistics*, 3(3), Article 100167. <https://doi.org/10.1016/j.rmal.2024.100167>
- Papi, M. (2010). The L2 motivational self system, L2 anxiety, and motivated behavior: A structural equation modeling approach. *System*, 38(3), 467–479. <https://doi.org/10.1016/j.system.2010.06.011>
- Pawlak, M., Kruk, M., Zawodniak, J., & Pasikowski, S. (2020). Investigating factors responsible for boredom in English classes: The case of advanced learners. *System*, 91, Article 102259. <https://doi.org/10.1016/j.system.2020.102259>

- Pawlak, M., Zarrinabadi, N., & Kruk, M. (2024). Positive and negative emotions, L2 grit and perceived competence as predictors of L2 motivated behaviour. *Journal of Multilingual and Multicultural Development*, 45(8), 3188–3204. <https://doi.org/10.1080/01434632.2022.2091579>
- Piechurska-Kuciel, E. (2011). Perceived teacher support and language anxiety in Polish secondary school EFL learners. *Studies in Second Language Learning and Teaching*, 1(1), 83–100. <https://doi.org/10.14746/ssllt.2011.1.1.5>
- Plonsky, L., & Gonulal, T. (2015). Methodological synthesis in quantitative L2 research: A review of reviews and a case study of exploratory factor analysis. *Language Learning*, 65(S1), 9–36. <https://doi.org/10.1111/lang.12111>
- Ryan, R. M., Connell, J. P., & Plant, R. W. (1990). Emotions in nondirected text learning. *Learning and Individual Differences*, 2(1), 1–17. [https://doi.org/10.1016/1041-6080\(90\)90014-8](https://doi.org/10.1016/1041-6080(90)90014-8)
- Saito, K., Dewaele, J.-M., Abe, M., & In'nami, Y. (2018). Motivation, emotion, learning experience, and second language comprehensibility development in classroom settings: A cross-sectional and longitudinal study. *Language Learning*, 68(3), 709–743. <https://doi.org/10.1111/lang.12297>
- Seligman, M. E. P. (2011). *Flourish: A visionary new understanding of happiness and well-being*. Free Press.
- Shimizu, H., Hoshino, Y., & Hamada, A. (2026). Multi-construct argument-based validation of foreign language emotion scales for young language learners. *Research Methods in Applied Linguistics*, 5(2), Article 100311. <https://doi.org/10.1016/j.rmal.2026.100311>
- Teimouri, Y., Goetze, J., & Plonsky, L. (2019). Second language anxiety and achievement: A meta-analysis. *Studies in Second Language Acquisition*, 41(2), 363–387. <https://doi.org/10.1017/s0272263118000311>
- Xethakis, L. (2024). A confirmatory and exploratory approach to the dimensions and sources of foreign language enjoyment among Japanese university English learners. *Journal of the School of Humanities and Science*, 2, 39–52.
- Xethakis, L., Rupp, M., Edwards, O., Howarth, M., & Kawagoe, T. (2025). Validating the Japanese version of the short-form foreign language enjoyment scale. *Research Methods in Applied Linguistics*, 4(3), Article 100270. <https://doi.org/10.1016/j.rmal.2025.100270>
- Yan, J. X., & Horwitz, E. K. (2008). Learners' perceptions of how anxiety interacts with personal and instructional factors to influence their achievement in English: A qualitative analysis of EFL learners in China. *Language Learning*, 58(1), 151–183. <https://doi.org/10.1111/j.1467-9922.2007.00437.x>
- Yang, L., & Lin, Y. (2026). To talk or to remain silent? Foreign language enjoyment and perceived classroom climate as drivers of change in Chinese EFL learners' willingness to communicate: A latent growth curve analysis. *Innovation in Language Learning and Teaching*, 20(4), 694–711. <https://doi.org/10.1080/17501229.2024.2391376>

APPENDICES

APPENDIX A.

The item description based on Dewaele and MacIntyre (2016)

- 1 I can be creative.
 - 2 I can laugh off embarrassing mistakes in the FL.
 - 3 I don't get bored.
 - 4 I enjoy it.
 - 5 I feel as though I'm a different person during the FL class.
 - 6 I learnt to express myself better in the FL.
 - 7 I'm a worthy member of the FL class.
 - 8 I've learnt interesting things.
 - 9 In class, I feel proud of my accomplishments.
 - 10 It's a positive environment.
 - 11 It's cool to know a FL.
 - 12 It's fun.
 - 13 Making errors is part of the learning process.
 - 14 The peers are nice.
 - 15 The teacher is encouraging.
 - 16 The teacher is friendly.
 - 17 The teacher is supportive.
 - 18 There is a good atmosphere.
 - 19 We form a tight group.
 - 20 We have common 'legends', such as running jokes.
 - 21 We laugh a lot.
-

APPENDIX B.

Items	FLES developed by Botes et al. (2021)	The Japanese version of the FLES
4	I enjoy it.	私は授業で、英語を学ぶのが楽しい。
8	I've learnt interesting things.	私は英語の授業では、興味深いことを学んでいる。
9	In class, I feel proud of my accomplishments.	私は英語の授業では、自分の学びの成果を誇りに思う。
15	The teacher is encouraging.	英語の先生はいつも励まそうとしてくれる。
16	The teacher is friendly.	英語の先生は、フレンドリー（親しみやすい）に接しようとしてくれる。
17	The teacher is supportive.	英語の先生は、サポート（支援）しようとしてくれる。
19	We form a tight group.	私たちの英語のクラスには、一体感がある。
20*	We have common 'legends,' such as running jokes.	私たちのクラスには、共通語のような内輪ネタなどがある。
21	We laugh a lot.	私たちは、英語の授業中によく笑いあう。

Note. * Item 20 was removed from the refined model.