



OPEN ACCESS

EDITED BY

Raona Williams,
Ministry of Education, United Arab Emirates

REVIEWED BY

Walter Capa-Luque,
National University Federico Villareal, Peru
Ioannis G. Katsantonis,
University of Cambridge, United Kingdom
Dias Tiara Putri Utomo,
Universitas Muhammadiyah Lamongan,
Indonesia
Abdullah Althamir,
University of Anbar, Iraq

*CORRESPONDENCE

Nataly Susan Saez-Zevallos
✉ natalysaez@upeu.edu.pe
Danitza Elfi Montalvo-Apolín
✉ danitza.montalvo@upeu.edu.pe

RECEIVED 16 January 2025

ACCEPTED 24 April 2025

PUBLISHED 13 May 2025

CITATION

Saez-Zevallos NS, Cunza-Aranzábal DF,
Abanto-Ramírez CD and
Montalvo-Apolín DE (2025) Psychometric
properties of the strategy inventory for
language learning in Peruvian university
students.

Front. Educ. 10:1562101.

doi: 10.3389/feduc.2025.1562101

COPYRIGHT

© 2025 Saez-Zevallos, Cunza-Aranzábal,
Abanto-Ramírez and Montalvo-Apolín. This is
an open-access article distributed under the
terms of the [Creative Commons Attribution
License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is permitted,
provided the original author(s) and the
copyright owner(s) are credited and that the
original publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Psychometric properties of the strategy inventory for language learning in Peruvian university students

Nataly Susan Saez-Zevallos^{1*}, Denis Frank Cunza-Aranzábal²,
Carlos D. Abanto-Ramírez¹ and Danitza Elfi Montalvo-Apolín^{1*}

¹Unidad de Postgrado de Ciencias Humanas y de la Educación, Universidad Peruana Unión, Lima, Peru, ²Facultad de Ciencias de la Salud, Universidad Peruana Unión, Tarapoto, Peru

Background: In contemporary academia, proficiency in a second language, particularly English, is of significant importance for university students. The objective of this research is to determine the validity and reliability properties of the Strategy Inventory for Language Learning (SILL) among Peruvian university students.

Methods: Six hundred ten Peruvian university students, aged between 18 and 50 years ($M = 22.8$; $SD = 5.90$), participated in the study. Statistical techniques for latent variables were employed, and recommendations for determining the psychometric properties of educational and psychological tests were followed.

Results: The results show that the SILL exhibits an internal structure organized into six first-order factors and one second-order factor, consistent with the theoretical proposal. This structure was confirmed through CFA with excellent goodness-of-fit indices ($\chi^2 = 222.254$, $df = 129$, $p = 0.000$, $RMSEA = 0.034$, $SRMR = 0.037$, $CFI = 0.971$, $TLI = 0.966$). The inventory also demonstrates adequate internal consistency for its six factors, as well as for the complete inventory ($\alpha/\omega/H > 0.7$), and is invariant according with gender (male and female), geographical location (coast, mountains and jungle) age (late adolescence and early adulthood), and English educational level (A and B).

Conclusion: It's concluded that the SILL is a measurement instrument that demonstrates validity, factorial invariance, and high reliability in its scores. Consequently, it can be utilized in future academic research.

KEYWORDS

language learning strategies, English language, university students, validity, reliability

1 Introduction

English language proficiency is currently recognized as a crucial competence for university students, as it facilitates greater access to information in scientific publications (Sri Andayani, 2022). This proficiency contributes to the enhancement of students' research skills, which are essential for their academic success (Karim et al., 2023), particularly in developing critical thinking, problem-solving abilities, affective responses (Vieno et al., 2022), and in achieving superior academic performance (Zhai and Razali, 2023). Furthermore, English language proficiency in the university context affords students improved academic and professional opportunities abroad, facilitating the process of cultural adaptation that such experiences entail (Wilczewski and Alon, 2023). Moreover, the significance of English proficiency extends

beyond the university environment, as it is identified as a key factor for the professional growth of graduates, providing enhanced employment opportunities (John et al., 2021).

In Peru, since the year 2000, the state has implemented English courses at all educational levels (British Council Perú, 2024). However, according to the latest English Proficiency Index (EPI, 2023) ranking, Peru occupies the 50th position and demonstrates a medium level of English proficiency. Considering these reports, it is of paramount importance to explore the English language learning strategies employed by students in order to implement pedagogical actions appropriate to their needs and learning styles (Liu et al., 2023).

Language learning strategies are defined as actions employed by students to enhance their proficiency in second language acquisition skills (Oxford, 1986, 1990). The English language learning strategies were theoretically classified into direct and indirect strategies (Oxford, 1990). Among the direct strategies are memory strategies which enable students to store and retrieve knowledge, employing techniques to organize, review, and retain linguistic material in memory (Ardasheva and Tretter, 2013; García-Herrero and Jiménez-Vivas, 2014; Wu and Chang, 2018) cognitive strategies involve the utilization of cognitive skills such as reasoning, analysis, and application to engage with linguistic material, facilitating more efficient language processing (Ardasheva and Tretter, 2013; García-Herrero and Jiménez-Vivas, 2014; Wu and Chang, 2018; John Peter and Hashim, 2023) and compensation strategies are employed to allow students to address deficiencies in their language knowledge, particularly when expressing themselves orally or in writing.

Concerning indirect strategies, metacognitive strategies encompass planning, monitoring, and self-evaluation of learning, enabling students to reflect on their progress and manage their mental processes autonomously (Ardasheva and Tretter, 2013; García-Herrero and Jiménez-Vivas, 2014; John Peter and Hashim, 2023) affective strategies facilitate students' management of their emotions, motivation, and attitudes towards learning, which are essential for maintaining interest and overcoming negative emotions such as anxiety or frustration (Ardasheva and Tretter, 2013; García-Herrero and Jiménez-Vivas, 2014; Wu and Chang, 2018; John Peter and Hashim, 2023). Social strategies involve interaction with others in the learning process (Ardasheva and Tretter, 2013; García-Herrero and Jiménez-Vivas, 2014; Wu and Chang, 2018; John Peter and Hashim, 2023). These strategies assist students in interacting with peers, instructors, and native speakers to enhance their linguistic skills.

Recent research has focused on enhancing English language acquisition in university settings. A study conducted on Chinese students demonstrated that their attitude towards internet usage influences online English learning, suggesting that educators and administrators should promote its utilization to enhance language proficiency (Wang et al., 2023). Furthermore, the implementation of artificial intelligence in English language instruction has been shown to improve competence, motivation, and self-regulation in language learning (Wei, 2023). Gamification has been identified as an effective strategy for English language acquisition, fostering an empathetic environment and enhancing relationships, motivation, and learning enjoyment (Laura-De La Cruz et al., 2023). The analysis of English-language news articles also facilitates language proficiency, strengthens social responsibility, and improves linguistic skills (Gong, 2023). International student exchange programs in universities contribute

significantly to English language proficiency and enrich academic development (Peng, 2021).

Language learning strategies employed by university students may influence their language proficiency (Kovacevic, 2019). The utilization of gamification enhances vocabulary acquisition and promotes motivation for collaborative work, although it is less effective for grammatical learning (Batlle Rodríguez and Argüello, 2023). Furthermore, implementing English-language conferences and utilizing English within the university environment, in conjunction with student motivation, would constitute effective strategies (Tai and Zhao, 2024). Thus, the strategies implemented by an educational institution can stimulate the use of learning strategies.

Although there are no studies on the combined effectiveness of various strategies in university students, it has been demonstrated that, in secondary school students, cognitive, metacognitive, affective, meta-affective, social, and meta-social strategies significantly influence attitudes and performance in English (Habók et al., 2022). A systematic review of 42 studies (2017–2023) reveals that, in Latin America, except for Chile, there is a lack of research on English learning strategies among university students, indicating limited exploration in this region (Domínguez and Juanías, 2024).

The extant literature identifies several scales for measuring this construct in university students. The Language Learner Factors Questionnaire, developed in China for tertiary students, evaluates factors influencing English learning through 17 items across three strategies: focus on form, focus on meaning, and avoidance of the mother tongue; however, validity and reliability evidence were not reported (Wen and Johnson, 1997). The Language Strategy Use Inventory (LSUI), developed in the United States, assesses learning strategies in university students with 90 items in six categories: listening, vocabulary, speaking, reading, writing, and translation, for which adequate reliability was reported (Cohen, 2014; Kappler Mikk et al., 2019).

The Strategy Inventory for Language Learning (SILL) is the most representative instrument for measuring English learning strategies, psychometrically analyzed in university contexts, albeit with a varying number of items. Unlike other scales that focus on a single type of strategy, the SILL comprehensively measures cognitive, metacognitive, affective, and social strategies (John Peter and Hashim, 2023). Upon its development, the scale was anticipated to identify six factors in accordance with the proposed theoretical framework—memory, cognitive, compensation, metacognitive, affective, and social strategies—however, the initial version of 135 items (SILL 1.3) revealed ten factors with reliability of the whole instrument ($\alpha = 0.9467$) and each of its factors. The reliability and validity of the internal structure of each factor was reported considering only 114 items for a second version; nevertheless, some factors exhibited a reliability below 0.7, and one factor did not demonstrate adequate reliability ($\alpha = 0.3061$). An alternative version consisting of 121 items was proposed (SILL 2.1); however, its reliability and validity were not reported. Additionally, another version comprising 47 items was proposed (SILL 2.2), but evidence of its reliability and validity was not provided (Oxford, 1986). Subsequently, the scale was modified and presented in two formats: one with 80 items (SILL 5.1) for English speakers learning another language and one with 50 items (SILL 7.0) for speakers of other languages learning English, both with a structure based on Oxford's system for classifying strategies into six groups, but evidence of factorial structure of this system was not reported (Oxford,

1990). This 50-item version was used in university students in Puerto Rico, yielding 9 factors: strategies for active and naturalistic use of English, metacognitive strategies with affective support, social and affective strategies, reflective strategies for language analysis and anxiety awareness, sensory memory strategies, cognitive and social strategies for conversation practice, sensory image strategies for vocabulary learning, strategies for cognitive manipulation of language, and general compensation strategies (Green and Oxford, 1995).

Psychometric analysis of the SILL was conducted in Taiwan, China, Japan, Egypt, and the United States showing 9 factors; however, only the reliability of the whole scale was reported (Oxford and Burry-Stock, 1995) as in Thailand (Zou and Lertlit, 2022). Nevertheless, in Japan, the internal structure of the 50-item SILL was evaluated through confirmatory factor analysis, revealing a lack of evidence for model fit (Isemonger, 2016). Meanwhile, in Brazil, a study examining participants from diverse educational backgrounds, including undergraduates, reported evidence of reliability for each factor. However, the affective strategies factor demonstrated a reliability coefficient of 0.566, while the memory factor exhibited a reliability of 0.691. Nevertheless, evidence of validity regarding the internal structure was not reported (Yang et al., 2021). A 38-item version was adapted for utilization with Spanish university students, demonstrating only the reliability of the complete instrument but not of its individual factors, without reporting evidence of validity evaluated in the assessed population (García-Herrero and Jiménez-Vivas, 2014). In Chile reliability of each factor was reported but not evidence of validity of the internal structure of 50 items grouped in nine factors (Barrios et al., 2017; Montaña-González and Cancino, 2020). Based on the aforementioned considerations, it is evident that the SILL demonstrates a theoretical framework with a wide use; however, it necessitates an assessment of its internal structure validity through confirmatory factor analysis, an evaluation of its construct validity, and a reliability analysis of its factors, particularly in the Spanish version.

According to Oxford (1990) the use of language learning strategies is a common phenomenon, with women typically using a broader and more diverse range of strategies compared to men (Sumarni and Rachmawaty, 2019). Both female and male often use social strategies while the lowest frequency strategy uses both were different. Ranjan and Philominraj (2020) noted that women, although with a not very significant difference, reported using language learning strategies slightly more frequently than men. In another recent study, between 1989 and 2021, women, in their highest percentage, respond significantly to Language Learning Strategies but no significant differences were found in productive vocabulary (Montero-SaizAja, 2021). In another study females use all the strategies more often than males particularly on metacognitive strategy (Andini and Prastiyowati, 2021). In consideration of the geographical context, it has been reported that the teaching and learning of the English language differ according to the geographical location of the student. This variation is influenced by factors such as socioeconomic and cultural status, linguistic and educational context, and the traditions of educators. These factors affect language proficiency, and the learning strategies employed (Hu, 2003; Habók et al., 2021). According to the language learning strategies (SILL) in A1–B1 level students, social strategies were the most frequently used, highlighting a preference for collaborative and interactive learning. Affective and metacognitive strategies, were more common at advanced levels (Gómez et al., 2021). According to the age, secondary (aged 13–18) and tertiary

(aged 20–22) students reported using compensation strategies more frequently than primary (aged 10–12) students. Tertiary students used social and affective strategies more frequently than did other age groups (Chen, 2014). Nonetheless, these investigations lacked a proper demonstration of measuring invariance. Measurement invariance constitutes a crucial prerequisite when comparing outcomes across distinct groups (e.g., men and women, different geographical location, age and educational level) to ensure that the instrument measures the same construct equivalently for all groups. Without establishing measurement invariance, group differences in scores may reflect measurement bias rather than true differences in the construct being measured (Vandenberg and Lance, 2000; Cheung and Rensvold, 2009).

The SILL in Spanish language, while demonstrating adequate global reliability, requires evaluation of its psychometric properties of internal structure through factor analysis, construct validity and the assessment of reliability of its factors. This aspect constitutes a research gap that warrants attention to ensure the precision of measurements to be conducted in future studies (Boateng et al., 2018), at the university education level. Consequently, the research question proposed in this study were: What are the psychometric properties reliability, validity of the internal structure, convergent validity, discriminant validity, and measurement invariance between genders (male and female), geographical location (coast, mountains and jungle), age (late adolescence and early adulthood), and English educational level (A and B) according to the common European framework (Council of Europe, 2001)? Five hypothesis correspondent to that question are proposed as follows:

- The SILL will demonstrate high internal consistency (α , ω and $H > 0.70$) when administered to non-native English speakers in a Peruvian university setting.
- The SILL will demonstrate adequate fit indices as evidence of validity of internal structure ($CFI/TLI > 0.90$ or 0.95 , $RMSEA < 0.05$ or 0.08 and $SRMR < 0.08$).
- The SILL will demonstrate evidence of convergent validity (AVE values ≥ 0.5).
- The SILL will demonstrate evidence of discriminant validity ($AVE > \rho^2$).
- The SILL will demonstrate evidence of measurement invariance for men and women, geographical location (coast, mountains and jungle) age (late adolescence and early adulthood), and English educational level (A and B) at the levels of configural, metric, scalar and strict invariance ($|\Delta CFI| < 0.01$, $|\Delta RMSEA| < 0.015$, and $|\Delta SRMR| < 0.03$).

Therefore, the objective of this study is to determine the validity and reliability properties of the Strategy Inventory for Language Learning (SILL) in Spanish with Peruvian university students. This instrument enables them to identify and implement effective practices for learning this language, thereby optimizing the teaching-learning process.

2 Method

2.1 Design

This study is categorized as quantitative, employing a survey-based, cross-sectional research design (Creswell and Creswell, 2023),

and is characterized as a psychometric study as it focuses on the validation and evaluation of a measurement instrument, to ensure the reliability and validity of its measurements for their intended purpose (Rust et al., 2021).

2.2 Participants

To determine the sample size, a sample size calculator for structural equation models was utilized (Soper, 2024), with an anticipated effect size of 0.3, a desired statistical power of 0.8, 6 latent variables, 38 observed variables, and a probability level of 0.05, which indicated that the recommended minimum size was 256. The data were collected from 610 university students from three branches of one university, located in the three geographical regions of Peru (coast, highlands and jungle), aged between 18 and 50 years ($M = 22.8$; $SD = 5.90$) of whom 399 (65.4%) were in the early adulthood (20 to 39 years) and 355 (58.2%) were female. Regarding marital status, 560 (91.8%) were single. With respect to faculty affiliation, 261 (42.8%) were from Health Sciences; and concerning geographical area, 342 (56.1%) were from the coastal region (Table 1).

2.3 Ethical aspects

To conduct this study, the processes outlined in the Declaration of Helsinki were followed, and approval was obtained from the ethics

TABLE 1 Sample description.

Variables	Categories	Frequency	%
Age (Min = 18; Max = 50; $M = 22.8$; $SD = 5.9$)	Late adolescence	194	31.8%
	Early adulthood	399	65.4%
	Middle adulthood	17	2.8%
Gender	Female	355	58.2%
	Male	255	41.8%
Marital status	Single	560	91.8%
	Married	44	7.2%
	Divorced	5	0.8%
	Widow(er)	1	0.2%
Faculty	Health Sciences	261	42.8%
	Business Studies	166	27.2%
	Engineering and Architecture	139	22.8%
	Theology	16	2.6%
	Human Sciences and Education	28	4.6%
English educational level	A	291	47.7%
	B	295	48.4%
	C	24	3.9%
Region	Coast	342	56.1%
	Mountains	155	25.4%
	Jungle	113	18.5%

committee of the researchers' affiliated university (Reference 2023-CEEPG-00010). Subsequently, the directors of each professional school and the coordinators of the university language center were contacted, explaining the nature of the research. They were requested to provide class schedules for virtual or in-person modalities, as well as the telephone numbers of the instructors to coordinate the date and time for the administration of electronic forms intended for data collection. Then, on the designated date and time, the link was sent via WhatsApp for in-person classes, while for virtual classes, the link was sent via Zoom chat. The electronic form, developed using Google Forms, included a format to obtain informed consent from each participant, followed by a section for collecting sociodemographic data and another section that included the data collection instrument.

2.4 Instruments

The Strategy Inventory for Language Learning (SILL) is an instrument developed by Oxford (1986, 1990), translated and adapted to Spanish for use with Spanish university students in a reduced form of 38 items (García-Herrero and Jiménez-Vivas, 2014) to ascertain the type of learning strategies employed by a student when learning the English language. It comprises six dimensions that measure the subjects' use of memory strategies (Items 1 to 8), cognitive strategies (Items 9 to 18), compensation strategies (Items 19 to 22), metacognitive strategies (Items 23 to 27), affective strategies (Items 28 to 31), and social strategies (Items 32 to 38), utilizing a 5-point Likert scale: "never or almost never" = 1, "generally not" (less than half the time) = 2, "sometimes" (approximately half the time) = 3, "often" (more than half the time) = 4, and "always or almost always" = 5. The Spanish version utilized herein demonstrated evidence of reliability ($\alpha = 0.823$); however, its evidences of content and construct validity was solely supported by the English version of 50 items (Oxford, 1990; García-Herrero and Jiménez-Vivas, 2014).

2.5 Data analysis

For statistical analysis, the free software R 4.3.1 (R Core Team, 2023) was utilized through the R Studio 2023.06.1 + 524 interface, where Confirmatory Factor Analysis and reliability processes were conducted.

Initially, a preliminary psychometric analysis of the items was conducted using the Pearson product-moment discrimination index ($DI > 0.2$) of each item in relation to the complete scale without that item (Rust et al., 2021), as well as a reliability analysis evaluating whether the α coefficient decreases when each item is removed.

Subsequently, prior to initiating the CFA, the corresponding assumptions were verified. The skewness and kurtosis were confirmed to have values between -1 and $+1$, which fit the assumption of univariate normality (Finney and DiStefano, 2013), and to evaluate the multivariate normal distribution assumption, the expected Mardia's skewness of 0 and kurtosis of $p(p+2)$ for p variables were considered. Values below this expectation were interpreted as indicators of platykurtosis, while higher values were interpreted as indicators of leptokurtosis. Moreover, if at least one of these tests is statistically significant (p -value < 0.05), it is inferred that the underlying distribution is non-normal (Lewis, 2013; Cain et al., 2017).

Subsequently, utilizing the *lavaan* library in the RStudio interface, Confirmatory Factor Analysis (CFA) was conducted using the Maximum Likelihood estimator with Satorra-Bentler correction (MLM), which represents an appropriate alternative when analyzing items as continuous variables with non-normal distribution (Finney and DiStefano, 2013). For the evaluation of the adjustment models, the following indicators were considered: the chi-square test (χ^2), the confirmatory fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA), and the standardized root mean square residual (SRMR), with CFI and TLI values > 0.90 indicating an acceptable fit and > 0.95 indicating a good fit (Schumacker and Lomax, 2016). For RMSEA values below 0.05 indicate a good fit and below 0.08 indicate an acceptable fit, and for SRMR values below 0.08 indicate a good fit (Bandalos and Finney, 2019; Kline, 2023).

The reliability of the latent model was calculated using Cronbach's α coefficients (Cronbach, 1951) and McDonald's ω (McDonald, 2009) from the psych statistical package (Revelle, 2022), as well as the H index (Hanckock and Mueller, 2013), for which magnitudes > 0.70 are considered adequate (Domínguez-Lara, 2016). Factor loadings with values $\lambda > 0.50$ were deemed appropriate. Convergent validity was estimated using the average variance extracted (AVE), where values should at least 0.50 to indicate that the observed items or variables correspond adequately to their factor, while discriminant validity would be assumed to the extent that the AVE of each latent variable was greater than the square of the correlation between them (ϕ^2), thus indicating that the observed variables of one factor are sufficiently distinct from other factors (Hair et al., 2019). The analysis to calculate invariance was conducted using the *hemp* package (Desjardins and Bulut, 2018). Factorial invariance between males and females was verified, evaluating the more restricted CFAs (Chen, 2007). The first level, configural invariance (M1), evaluates the reference model; the second level, metric invariance (M2), assesses the equality of factor loadings; the third level, scalar invariance (M3), examines the equality of indicator intercepts; and the fourth level, strict invariance (M4), evaluates the equality of indicator residuals (Meredith, 1993; Brown, 2015). To evaluate the invariance of the factorial structure across the groups under comparison, models of first and second order were compared (Chen et al., 2005). As the chi-square difference ($\Delta\chi^2$ non-significant) is sensitive to sample size, the absolute value of the CFI difference ($|\Delta CFI| < 0.01$), RMSEA ($|\Delta RMSEA| < 0.015$), and SRMR ($|\Delta SRMR| < 0.03$) were utilized (Putnick and Bornstein, 2016; Finch and French, 2018).

3 Results

3.1 Preliminary analysis of the items

A descriptive analysis of the 38 items in the scale under evaluation was conducted, calculating the descriptive statistics of mean, median, maximum and minimum values, skewness (g_1), and kurtosis (g_2). Additionally, the discrimination index (DI) was evaluated using Pearson's product-moment correlation coefficient for each item in relation to the other items on the scale. The reliability (α) of the complete scale was analyzed if an item was eliminated (Table 2). All items met the assumption of univariate normality, as the skewness (g_1) and kurtosis (g_2) were within the range of ± 1 , the DI values were > 0.2 ,

and the α value of 0.955 for the complete scale decreased if any item was removed, with the exception of item 36. However, it was decided to retain item 36 due to the adequacy of its other descriptive measures. The evaluation of multivariate normal distribution through Mardia's test indicated that the multivariate normality assumption was not met, as the multivariate skewness was 22812.21 (expected value = 0) and kurtosis was 115.04 (expected value = 1,520), both with p -value < 0.05 .

3.2 Validity of the internal structure and reliability

In order to verify the validity of the internal structure of the instrument, confirmatory factor analysis (CFA) was carried out, obtaining adequate and adequate fit indices ($\chi^2 = 1428.958$, $df = 650$, $p = 0.000$, RMSEA = 0.044, SRMR = 0.049, CFI = 0.905, TLI = 0.897) with adequate evidence of reliability (α , ω , $H > 0.7$). However, there was insufficient evidence of convergent validity (AVE < 0.5) and lack evidence of discriminant validity ($\phi^2 > AVE$) as be seen in Table 3.

The absence of convergent and discriminant validity evidence for the 38-item, six-factor model prompted the hypothesis that a more robust model could be developed. This enhanced framework would be achieved through strategic item reduction, retaining only those elements that contribute to a more stable structure and exhibit clear convergent and discriminant validity properties. This second model of 20 items in six factors was evaluated using CFA, and adequate fit indices were obtained ($\chi^2 = 181.570$, $df = 120$, $p = 0.000$, RMSEA = 0.029, SRMR = 0.032, CFI = 0.981, TLI = 0.976), with evidence of convergent validity at one decimal place (AVE = 0.5), lack evidence of discriminant validity ($\phi^2 > AVE$), and evidence of reliability (α , ω , $H > 0.7$).

3.3 Second order model

As favorable fit indices were obtained for a six-factor model, but evidence of high correlations between factors (as shown in Table 3), it was hypothesized that a model with six first-order factors and one second-order factor encompassing the first-order factors might exist. This second model (Figure 1) was evaluated using CFA, and adequate fit indices were also obtained ($\chi^2 = 222.254$, $df = 129$, $p = 0.000$, RMSEA = 0.034, SRMR = 0.037, CFI = 0.971, TLI = 0.966), although not superior to the indices of the first evaluated model. Nevertheless, it is considered that this model would better explain the high correlation between some primary factors ($\phi > 0.7$) based on the notion of all items belonging to a second-order factor. The reliability of this second order model was adequate ($\alpha = 0.916$; $\omega = 0.917$).

3.4 Measurement invariance analysis

The second-order model obtained was assessed for its invariance across the entire sample ($N = 610$), taking into account gender (males and females) and geographical location (coast, mountains, and jungle). To address potential extremal differences between groups, the sample was reduced based on age (late adolescence and early adulthood) and English educational level (A and B), resulting in sample sizes of 593 and 586, respectively.

TABLE 2 Descriptive analysis of the items.

	Mean	SD	Min.	Max.	g_1	g_2	DI	α
Item1	3.61	0.87	1	5	−0.240	0.027	0.532	0.954
Item2	3.48	0.88	1	5	−0.254	0.049	0.621	0.953
Item3	3.64	0.95	1	5	−0.464	0.061	0.496	0.954
Item4	3.33	1.03	1	5	−0.229	−0.284	0.535	0.954
Item5	3.05	1.12	1	5	−0.058	−0.708	0.534	0.954
Item6	3.55	0.92	1	5	−0.394	0.027	0.584	0.954
Item7	3.47	0.91	1	5	−0.177	−0.247	0.701	0.953
Item8	3.50	0.93	1	5	−0.264	−0.282	0.572	0.954
Item9	3.59	0.92	1	5	−0.333	−0.120	0.575	0.954
Item10	3.73	0.91	1	5	−0.240	−0.480	0.538	0.954
Item11	3.73	0.87	1	5	−0.392	0.092	0.598	0.953
Item12	3.37	0.98	1	5	−0.131	−0.227	0.592	0.953
Item13	3.40	0.91	1	5	−0.061	−0.082	0.645	0.953
Item14	3.75	0.92	1	5	−0.396	−0.172	0.595	0.953
Item15	3.68	0.87	1	5	−0.145	−0.374	0.666	0.953
Item16	3.80	0.95	1	5	−0.432	−0.314	0.578	0.954
Item17	3.51	0.95	1	5	−0.262	−0.222	0.674	0.953
Item18	3.79	0.84	1	5	−0.268	−0.292	0.698	0.953
Item19	3.72	0.88	1	5	−0.340	−0.028	0.550	0.954
Item20	3.63	0.81	1	5	−0.084	−0.134	0.642	0.953
Item21	3.42	0.94	1	5	−0.296	−0.155	0.581	0.954
Item22	3.53	0.91	1	5	−0.156	−0.204	0.605	0.953
Item23	3.88	0.88	1	5	−0.439	−0.221	0.593	0.953
Item24	3.83	0.76	1	5	−0.229	−0.190	0.665	0.953
Item25	3.78	0.81	1	5	−0.245	−0.254	0.679	0.953
Item26	3.40	0.94	1	5	−0.124	−0.306	0.620	0.953
Item27	3.83	0.82	1	5	−0.356	−0.072	0.605	0.953
Item28	3.63	0.89	1	5	−0.432	0.178	0.556	0.954
Item29	3.76	0.96	1	5	−0.511	−0.013	0.556	0.954
Item30	4.04	0.85	1	5	−0.499	−0.392	0.565	0.954
Item31	3.77	0.88	1	5	−0.275	−0.314	0.626	0.953
Item32	3.59	0.91	1	5	−0.261	−0.153	0.460	0.954
Item33	3.41	0.94	1	5	−0.287	0.080	0.596	0.953
Item34	3.46	0.97	1	5	−0.376	0.077	0.504	0.954
Item35	3.44	0.96	1	5	−0.328	−0.177	0.619	0.953
Item36	3.34	1.04	1	5	−0.221	−0.417	0.363	0.955
Item37	3.40	0.89	1	5	−0.012	−0.132	0.561	0.954
Item38	3.55	0.97	1	5	−0.137	−0.519	0.638	0.953

g_1 = skewness; g_2 = Kurtosis; DI = Discrimination index, α = Cronbach's alpha if the item is dropped.

This reduction was necessary due to the limited number of respondents in the third age category (middle adulthood: $n = 17$) and the third English educational level ($C = 24$). Invariance was confirmed using the ΔCFI , $\Delta RMSEA$, and $\Delta SRMR$ criteria. The imposition of the constraint of equal means did not significantly affect the model fit, indicating that the latent measures are comparable across both genders, the three geographical locations, the two age stages, and the two English educational levels. Consequently, the proposed model, unrestricted M1, at the first order M2, M4, M6, and at the second order M3, M5, and M7, satisfies the expected criteria and confirms the factorial invariance of the SILL (Table 4).

TABLE 3 Confirmatory factor analysis with evidence of reliability and construct validity.

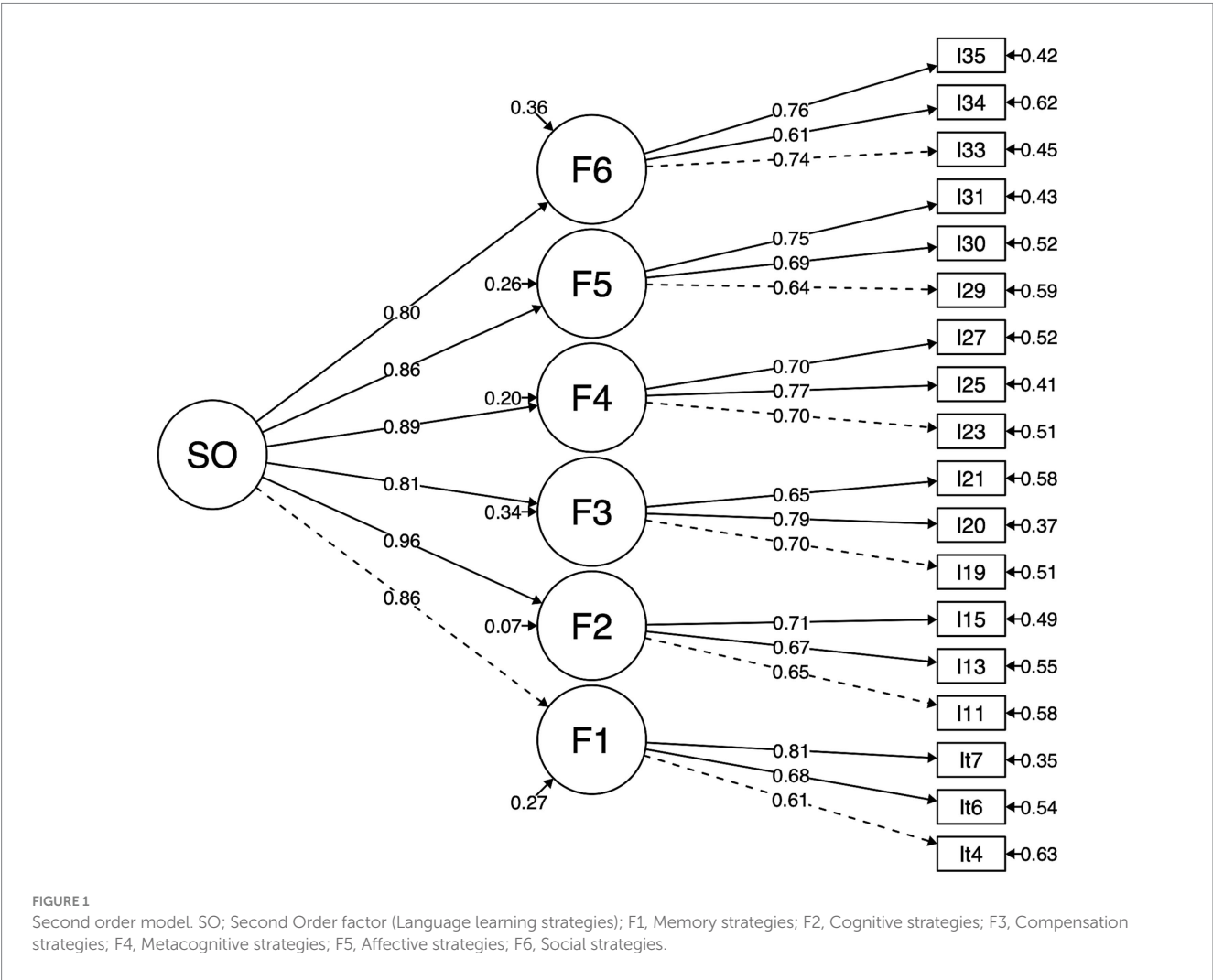
Initial number	CFA-1						CFA-2						Final number
	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	
Item1	0.60						-						-
Item2	0.71						-						-
Item3	0.57						-						-
Item4	0.62						0.61						Item1
Item5	0.65						-						-
Item6	0.66						0.67						Item2
Item7	0.77						0.82						Item3
Item8	0.61						-						-
Item9		0.59						-					-
Item10		0.60						-					-
Item11		0.66						0.64					Item4
Item12		0.62						-					-
Item13		0.66						0.67					Item5
Item14		0.65						-					-
Item15		0.73						0.72					Item6
Item16		0.63						-					-
Item17		0.71						-					-
Item18		0.75						-					-
Item19			0.65						0.69				Item7
Item20			0.74						0.80				Item8
Item21			0.66						0.65				Item9
Item22			0.63						-				-
Item23				0.70						0.69			Item10
Item24				0.78						-			-
Item25				0.78						0.76			Item11
Item26				0.64						-			-
Item27				0.70						0.71			Item12
Item28					0.61						-		-
Item29					0.65						0.64		Item13
Item30					0.70						0.71		Item14
Item31					0.70						0.74		Item15
Item32						0.50						-	-
Item33						0.70						0.75	Item16
Item34						0.58						0.61	Item17
Item35						0.73						0.76	Item18
Item36						0.46						-	-
Item37						0.64						-	-
Item38						0.69						-	-
α	0.85	0.88	0.76	0.84	0.76	0.81	0.73	0.72	0.75	0.76	0.74	0.75	α
ω	0.85	0.88	0.76	0.84	0.76	0.81	0.74	0.72	0.75	0.76	0.74	0.75	ω
H	0.96	0.98	0.88	0.93	0.88	0.95	0.81	0.81	0.82	0.82	0.81	0.82	H
AVE	0.42	0.44	0.45	0.52	0.44	0.39	0.49	0.46	0.51	0.52	0.49	0.50	AVE
F1	1.00	0.73	0.64	0.54	0.50	0.56	1.00	0.78	0.46	0.53	0.45	0.51	F1

(Continued)

TABLE 3 (Continued)

Initial number	CFA-1						CFA-2						Final number
	F1	F2	F3	F4	F5	F6	F1	F2	F3	F4	F5	F6	
F2	0.85	1.00	0.73	0.73	0.69	0.62	0.88	1.00	0.69	0.67	0.64	0.54	F2
F3	0.80	0.86	1.00	0.63	0.54	0.67	0.68	0.83	1.00	0.52	0.41	0.40	F3
F4	0.73	0.86	0.80	1.00	0.79	0.58	0.73	0.82	0.72	1.00	0.76	0.50	F4
F5	0.71	0.83	0.74	0.89	1.00	0.64	0.67	0.80	0.64	0.87	1.00	0.52	F5
F6	0.75	0.79	0.82	0.76	0.80	1.00	0.71	0.73	0.63	0.71	0.72	1.00	F6

F1, Memory strategies; F2, Cognitive strategies; F3, Compensation strategies; F4, Metacognitive strategies; F5, Affective strategies; F6, Social strategies; λ , Factor loading; α = Cronbach's alpha; ω = McDonald's Omega; H = index H of reliability; AVE, average variance extracted; below the diagonal, inter-factor correlations; above the diagonal, variance shared between factors (ϕ^2).



4 Discussion

English language learning strategies among university students encompass the techniques and methods employed by students to facilitate the acquisition and mastery of English as a foreign language. In Peru and other Latin American countries, there is a need for appropriate instruments to measure English language learning strategies, considering dimensions such as

memory, cognitive, compensation, metacognitive, affective, and social strategies. The objective of this study was to determine the psychometric properties of the Strategy Inventory for Language Learning (SILL) and provide evidence of its validity, reliability, and invariance according to sex in Peruvian university students.

Regarding the evidence of validity of the internal structure of the SILL in Spanish language, the present study, through

TABLE 4 Analysis of invariance.

Model	χ^2 (df)	CFI	RMSEA [90% CI]	SRMR	Comp. model	$\Delta\chi^2$ (Δ df)	Δ CFI	Δ RMSEA	Δ SRMR	Decision
Gender										
M1: Configural	538.515 (258)	0.938	0.060 [0.053; 0.067]	0.045						
M2: Metric 1st order	564.331 (275)	0.936	0.059 [0.052; 0.066]	0.055	M1	25.816 (17)	−0.002	−0.001	0.011	Accept
M3: Metric 2nd order	571.654 (282)	0.936	0.058 [0.051; 0.065]	0.058	M2	7.323 (7)	0.000	−0.001	0.003	Accept
M4: Scalar 1st order	587.036 (286)	0.933	0.059 [0.052; 0.066]	0.056	M3	15.382 ** (4)	−0.003	0.001	−0.002	Accept
M5: Scalar 2nd order	594.371 (293)	0.933	0.058 [0.051; 0.065]	0.059	M4	7.335 (7)	0.000	−0.001	0.003	Accept
M6: Residual 1st order	611.508 (304)	0.932	0.058 [0.051; 0.064]	0.056	M5	17.137 (11)	−0.001	0.000	−0.003	Accept
M7: Residual 2nd order	617.095 (311)	0.932	0.057 [0.050; 0.063]	0.058	M6	5.586 (7)	0.000	−0.001	0.002	Accept
Geographical location										
M1: Configural	775.496 (387)	0.916	0.070 [0.063; 0.077]	0.050						
M2: Metric 1st order	808.059 (421)	0.916	0.067 [0.060; 0.074]	0.062	M1	32.563 (34)	0.000	−0.003	0.012	Accept
M3: Metric 2nd order	836.83 (435)	0.913	0.067 [0.061; 0.074]	0.069	M2	28.771* (14)	−0.003	0.000	0.007	Accept
M4: Scalar 1st order	838.624 (443)	0.914	0.066 [0.059; 0.073]	0.063	M3	1.794 (8)	0.001	−0.001	−0.006	Accept
M5: Scalar 2nd order	867.47 (457)	0.911	0.066 [0.060; 0.073]	0.070	M4	28.846* (14)	−0.003	0.000	0.007	Accept
M6: Residual 1st order	918.247 (479)	0.905	0.067 [0.061; 0.074]	0.065	M5	50.777*** (22)	−0.006	0.001	−0.005	Accept
M7: Residual 2nd order	973.757 (493)	0.896	0.069 [0.063; 0.076]	0.076	M6	55.510*** (14)	−0.009	0.002	0.011	Accept
Age										
M1: Configural	595.609 (258)	0.923	0.066 [0.059; 0.073]	0.046						
M2: Metric 1st order	606.513 (275)	0.925	0.064 [0.057; 0.071]	0.050	M1	10.904 (17)	0.002	−0.002	0.004	Accept
M3: Metric 2nd order	622.257 (282)	0.923	0.064 [0.057; 0.071]	0.051	M2	15.744* (7)	−0.002	0.000	0.001	Accept
M4: Scalar 1st order	616.303 (286)	0.925	0.062 [0.056; 0.069]	0.051	M3	−5.954 (4)	0.002	−0.002	0.000	Accept
M5: Scalar 2nd order	632.18 (293)	0.923	0.062 [0.056; 0.069]	0.051	M4	15.877* (7)	−0.002	0.000	0.000	Accept
M6: Residual 1st order	642.924 (304)	0.923	0.061 [0.055; 0.068]	0.052	M5	10.744 (11)	0.000	−0.001	0.001	Accept
M7: Residual 2nd order	659.503 (311)	0.921	0.061 [0.055; 0.068]	0.053	M6	16.579* (7)	−0.002	0.000	0.001	Accept
English educational level										
M1: Configural	502.961 (258)	0.943	0.057 [0.049; 0.064]	0.044						
M2: Metric 1st order	523.079 (275)	0.942	0.055 [0.048; 0.063]	0.054	M1	20.118 (17)	−0.001	−0.002	0.010	Accept
M3: Metric 2nd order	541.066 (282)	0.940	0.056 [0.049; 0.063]	0.069	M2	17.987* (7)	−0.002	0.001	0.015	Accept
M4: Scalar 1st order	533.795 (286)	0.942	0.054 [0.047; 0.061]	0.054	M3	−7.271 (4)	0.002	−0.002	−0.015	Accept

(Continued)

TABLE 4 (Continued)

Model	χ^2 (df)	CFI	RMSEA [90% CI]	SRMR	Comp. model	$\Delta\chi^2$ (Δ df)	Δ CFI	Δ RMSEA	Δ SRMR	Decision
M5: Scalar 2nd order	551.807 (293)	0.940	0.055 [0.048; 0.062]	0.070	M4	18.012* (7)	-0.002	0.001	0.016	Accept
M6: Residual 1st order	546.565 (304)	0.943	0.052 [0.045; 0.059]	0.055	M5	-5.242 (11)	0.003	-0.003	-0.015	Accept
M7: Residual 2nd order	567.211 (311)	0.940	0.053 [0.046; 0.060]	0.070	M6	20.646** (7)	-0.003	0.001	0.015	Accept

Gender: N = 610, men (n = 255), women (n = 355); Region: N = 610, coast (n = 342), mountains (n = 155), jungle (n = 113); Age: N = 593, late adolescence (n = 194), early adulthood (n = 399); English educational level: N = 586, A (n = 291), B (n = 295). *p ≤ 0.05. **p ≤ 0.01. ***p ≤ 0.001.

Confirmatory Factor Analysis (CFA), assessed an initial model comprising 38 items grouped into six correlated factors. However, this model did not exhibit adequate Tucker-Lewis Index (TLI) values and lacked evidence of convergent and discriminant validity. Subsequently, a second model consisting of 18 items grouped into six correlated factors was evaluated to achieve convergent validity. Nevertheless, this second model demonstrated inadequate discriminant validity and high correlation between factors. Consequently, a third model—incorporating a second-order factor comprising the six factors of the second model—was evaluated, which provided evidence supporting a structure of one second-order factor with six first-order factors encompassing 18 items. This structure facilitates a comprehensive measurement of language learning strategies (cognitive, metacognitive, affective, and social) across the six *a priori* dimensions theoretically proposed in the original version and provides a whole scale score. The model demonstrates favorable fit indices through CFA (RMSEA, SRMR < 0.05; CFI, TLI > 0.95) as evidence of validity of internal structure, which is not observed in other six-dimensional versions, such as the Spanish version of 38 items (Green and Oxford, 1995; Oxford and Burry-Stock, 1995; García-Herrero and Jiménez-Vivas, 2014), the Thailand version of 50 items (Zou and Lertlit, 2022), or the 50-item Japan version where CFA was conducted with acceptable RMSEA and SRMSR but TLI and CFI indicated strong evidence against the model (RMSEA = 0.60, SRMR = 0.62; CFI = 0.772, TLI = 0.760) (Isemonger, 2016). Furthermore, the factor structure of six factors comprising 18 items presented in this study is more parsimonious and theoretically consistent than other versions of 50 items grouped into nine factors, such as the Spanish-language Puerto Rican and Chilean versions, which contain 50 items grouped into nine factors (Green and Oxford, 1995; Barrios et al., 2017; Montañó-González and Cancino, 2020) or other versions of nine factors psychometrically assessed in Taiwan, China, Japan, Egypt, and the United States (Oxford and Burry-Stock, 1995).

The reliability of the SILL was evaluated for the complete instrument, as well as its dimensions, using Cronbach's alpha coefficients (α), McDonald's Omega (ω), and coefficient H, whose values were greater than 0.7, indicating that the SILL is reliable, demonstrating precision when utilized as a measurement tool (Kline, 2023).

Measurement invariance has been confirmed across gender, geographical location, age and English educational levels. This validation supports unbiased comparisons between male and female students, from coastal, mountainous, and jungle regions, in the age stages of late adolescence and early adulthood, and individuals at A and B English educational levels according to the commonly educational framework, in future studies (Vandenberg and Lance, 2000; Cheung and Rensvold, 2009), with greater precision than the results of previous investigations that did not perform invariance prior to the comparison between male and female university students (Oxford, 1990; Sumarni and Rachmawaty, 2019; Ranjan and Philominraj, 2020; Andini and Prastiyowati, 2021; Montero-SaizAja, 2021), between students of different geographical location (Hu, 2003; Habók et al., 2021), between English levels according to the commonly educational framework (Gómez et al., 2021) and between students of different age groups (Chen, 2014).

4.1 Implications

English language learning strategies significantly influence the academic performance of university students and their ability to communicate in a foreign language, thereby facilitating proficiency in oral and written expression and comprehension in English. An instrument to assess these strategies can guide educators and administrators of educational institutions in generating a diagnosis and identifying areas where students may require additional support, with the aim of promoting the appropriate use of strategies that have a low level of utilization and, ultimately, improving language mastery. Furthermore, for students, it would enable the enhancement of their academic performance by allowing them to recognize their strengths and weaknesses, thus promoting the adoption or improvement of strategies that are underutilized or not employed in their English language learning, thereby fostering autonomous learning and the acquisition of greater confidence in the acquired language skills. By considering these implications, it is possible to create a learning environment that not only cultivates language proficiency but also develops critical skills for the academic and professional future of university students.

4.2 Limitations

While the current study established the reliability of the SILL through internal consistency and validated its internal structure, it did not obtain evidence of validity concerning other constructs or instruments measuring the same construct. Consequently, it is advisable for future research to undertake such analyses and to evaluate the psychometric properties of other instruments assessing English language learning strategies. Another limitation pertains to the evidence of invariance, which was demonstrated with respect to gender (male and female), geographical location (coast, mountains and jungle), age (late adolescence and early adulthood), and English educational level (A and B). However, it did not encompass all age stages (e.g., middle adulthood and late adulthood) or English educational levels (level C), as such analyses would require a larger sample size in one or more categories of these variables. Additionally, it is noteworthy that the data were collected virtually; therefore, it is recommended to perform further confirmatory factor analyses using data collected through printed formats and to compare those results with the findings of this study.

5 Conclusion

The SILL is a valid and reliable instrument that assesses six dimensions: memory strategies, cognitive strategies, compensation strategies, metacognitive strategies, affective strategies, and social strategies to measure English language learning strategies in Peruvian university students. Furthermore, this research has demonstrated that the scale is invariant across gender (male and female), geographical location (coast, mountains and jungle), age (late adolescence and early adulthood), and English educational

level (A and B), allowing for comparisons between the categories evaluated of those variables. This inventory contributes to the field of research in teaching and learning English as a second language in Spanish-speaking contexts. Researchers and practitioners seeking to understand and enhance English learning strategies and, ultimately, student performance and motivation, may find this scale to be a valuable tool. These findings have potential implications for educational practices and policies, with a focus on fostering English learning strategies and addressing differences in their use between males and females. Given its psychometric properties, it is recommended that the SILL also be validated in the context of regular basic education and at the postgraduate level.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Ethics statement

The studies involving humans were approved by Comité de ética de la Escuela de Posgrado de la Universidad Peruana Unión. The studies were conducted in accordance with the local legislation and institutional requirements. The participants provided their written informed consent to participate in this study.

Author contributions

NS-Z: Conceptualization, Data curation, Funding acquisition, Investigation, Methodology, Resources, Writing – original draft, Writing – review & editing. DC-A: Conceptualization, Data curation, Formal analysis, Methodology, Project administration, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. CA-R: Project administration, Supervision, Writing – original draft, Writing – review & editing. DM-A: Investigation, Project administration, Supervision, Writing – original draft, Writing – review & editing.

Funding

The author(s) declare that no financial support was received for the research and/or publication of this article.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The authors declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations,

or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/feduc.2025.1562101/full#supplementary-material>

References

- Andini, T. M., and Prastiyowati, S. (2021). Gender differences learning strategy at English language education department students university of Muhammadiyah Malang. *JINo* 7, 217–226. doi: 10.22219/jinop.v7i2.10476
- Ardasheva, Y., and Tretter, T. R. (2013). Strategy inventory for language learning–ELL student form: testing for factorial validity. *Mod. Lang. J.* 97, 474–489. doi: 10.1111/J.1540-4781.2013.12011.X
- Bandalos, D. L., and Finney, S. J. (2019). "Factor analysis: exploratory and confirmatory" in *The reviewer's guide to quantitative methods in the social sciences*. eds. G. R. Hancock, L. M. Stapleton and R. O. Mueller (New York, NY: Routledge), 98–120.
- Barrios, E., Montijano, M. P., Barrios, E., and Montijano, M. P. (2017). Self-reported strategy use and its variation according to proficiency among university learners of English as a foreign language. *Rev. Signos* 50, 312–336. doi: 10.4067/S0718-09342017000300312
- Batlle Rodríguez, J., and Argüello, V. G. (2023). Gamification and learning Spanish as a modern language: student perceptions in the university context. *Lang. Learn. High. Educ.* 13, 89–103. doi: 10.1515/CERCLIS-2023-2016/MACHINEREADABLECITATION/RIS
- Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quinonez, H. R., and Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Front. Public Health* 6:149. doi: 10.3389/fpubh.2018.00149
- British Council Perú (2024). Enseñanza del idioma inglés en Perú. Available at: <https://www.britishcouncil.pe/programas/educacion/ingles> (Accessed October 2, 2024).
- Brown, T. A. (2015). Confirmatory factor analysis for applied research. *Segunda Edn.* New York, NY, US: The Guilford Press.
- Cain, M. K., Zhang, Z., and Yuan, K. H. (2017). Univariate and multivariate skewness and kurtosis for measuring nonnormality: prevalence, influence and estimation. *Behav. Res. Methods* 49, 1716–1735. doi: 10.3758/s13428-016-0814-1
- Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct. Equ. Modeling* 14, 464–504. doi: 10.1080/10705510701301834
- Chen, M. L. (2014). Age differences in the use of language learning strategies. *Engl. Lang. Teach.* 7:p144. doi: 10.5539/ELT.V7N2P144
- Chen, F. F., Sousa, K. H., and West, S. G. (2005). Testing measurement invariance of second-order factor models. *Struct. Equ. Model.* 12, 471–492. doi: 10.1207/s15328007sem1203_7
- Cheung, G. W., and Rensvold, R. B. (2009). Evaluating goodness-of-fit indexes for testing measurement invariance. *Struct. Equ. Modeling* 9, 233–255. doi: 10.1207/S15328007SEM0902_5
- Cohen, A. D. (2014). Strategies in learning and using a second language. *2nd Edn.* New York, NY: Routledge.
- Council of Europe (2001). Common European framework of reference for languages: Learning, teaching, assessment. Cambridge: Cambridge University Press.
- Creswell, J. W., and Creswell, J. D. (2023). Research design: Qualitative, quantitative, and mixed methods approaches. *4th Edn.* Los Angeles: SAGE publications.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika* 16, 297–334. doi: 10.1007/BF02310555
- Desjardins, C. D., and Bulut, O. (2018). Handbook of educational measurement and psychometric using R. Boca Raton, FL: CRC Press, Taylor and Francis Group.
- Domínguez, J. M. M., and Juanías, J. M. (2024). Language learning strategies research in English as foreign language contexts: a systematic literature review. *J. Educ. Learn.* 18, 471–479. doi: 10.11591/EDULEARN.V18I2.21144
- Domínguez-Lara, S. A. (2016). Evaluación de la confiabilidad del constructo mediante el coeficiente H: breve revisión conceptual y aplicaciones. *Psychologia* 10, 87–94. doi: 10.21500/19002386.2134
- EPI (2023). The largest global ranking according to your English proficiency. Available at: <https://www.ef.com.co/epi/> (Accessed April 28, 2025).
- Finch, W. H., and French, B. F. (2018). A simulation investigation of the performance of invariance assessment using equivalence testing procedures. *Struct. Equ. Modeling* 25, 673–686. doi: 10.1080/10705511.2018.1431781
- Finney, S. J., and DiStefano, C. (2013). "Nonnormal and categorical data in structural equation modeling" in *Structural equation modeling: A second course* (Charlotte, NC: Information Age Publishing Inc.).
- García-Herrero, M. D. L. M., and Jiménez-Vivas, A. (2014). Estrategias en el aprendizaje de la lengua extranjera y niveles de competencia en estudiantes universitarios de magisterio. *Rev. Investig. Educ.* 32, 363–378. doi: 10.6018/RIE.32.2.167421
- Gómez, J. F., Díaz Larenas, C., and Gómez Torres, W. D. (2021). Estrategias de aprendizaje y creencias sobre el idioma inglés: una aproximación correlacional en estudiantes universitarios. *Praxis Educ.* 16, 1–17. doi: 10.5212/PraxEduc.v.16.16572.024
- Gong, W. (2023). Enhancing ESL learner's literacy by peer-assisted learning strategy of online English news. *Front. Psychol.* 14:1172099. doi: 10.3389/fpsyg.2023.1172099
- Green, J. M., and Oxford, R. (1995). A closer look at learning strategies, L2 proficiency, and gender. *TESOL Q.* 29:261. doi: 10.2307/3587625
- Habók, A., Kong, Y., Ragchaa, J., and Magyar, A. (2021). Cross-cultural differences in foreign language learning strategy preferences among Hungarian. *Chinese Mongolian Univ. Stud.* 7:e06505. doi: 10.1016/j.heliyon.2021.e06505
- Habók, A., Magyar, A., and Molnár, G. (2022). Investigating the relationship among English language learning strategies, language achievement, and attitude. *Front. Psychol.* 13:867714. doi: 10.3389/fpsyg.2022.867714
- Hair, J. F., Black, W. C., Babin, B. J., and Anderson, R. E. (2019). Multivariate data analysis. *8th Edn.* North Way: Cengage Learning.
- Hancock, G. R., and Mueller, R. O. (2013). *Structural equation modeling, a second course*, second Edn. Charlotte, NC: Information Age Publishing Inc.
- Hu, G. (2003). English language teaching in China: regional differences and contributing factors. *J. Multiling. Multicult. Dev.* 24, 290–318. doi: 10.1080/01434630308666503
- Isemonger, I. (2016). A confirmatory approach to the structural validity of scores generated by the strategy inventory for language learning (SILL). *Int. J. Soc. Cult. Stud.* 9, 1–14. (Accessed February 4, 2025).
- John, S., Gurario, S., and Halepota, J. A. (2021). The role of English language skills in career growth: a study of perceptions and strategies used to improve English language skills by graduate and undergraduate students of Karachi, Pakistan. *Global Soc. Sci. Rev.* VI, 346–355. doi: 10.31703/gssr.2021(VI-I).35
- John Peter, A. A., and Hashim, H. (2023). Language learning strategies used by ESL/EFL learners: a systematic review (2019 – 2023). *Int. J. Acad. Res. Bus. Soc. Sci.* 13:20148. doi: 10.6007/IJARBS/v13-i12/20148
- Kappler Mikk, B., Cohen, A. D., Paige, M., Chi, J. C., Lassegard, J. P., Maegher, M., et al. (2019). Maximizing study abroad: an instructional guide to strategies for language and culture learning and use. Available online at: <https://archive.carla.umn.edu/maxsa/guides.html#professionals> (Accessed December 16, 2024).
- Karim, A., Islam, M. S., Hamid, M. O., Rahman, M. M., and Amin, E. U. (2023). Language ideology, development of English proficiency, and performance in professional communication: voices of STEM + business graduates of English medium university. *Asia-Pac. J. Second Foreign Lang. Educ.* 8:27. doi: 10.1186/s40862-023-00200-w
- Kline, R. B. (2023). Principles and practice of structural equation modeling. *5th Edn.* New York, NY: The Guilford Press.
- Kovacevic, E. (2019). The relationship between language learning strategies and lexical complexity measures | las relaciones entre las estrategias de aprendizaje de idiomas y las medidas de complejidad léxica. *Porta Linguarum* 2019, 37–52. Available online at: <https://dialnet.unirioja.es/servlet/articulo?codigo=7186362> (Accessed February 17, 2025).

- Laura-De La Cruz, K. M., Noa-Copaja, S. J., Turpo-Gebera, O., Montesinos-Valencia, C. C., Bazán-Velasquez, S. M., and Pérez-Postigo, G. S. (2023). Use of gamification in English learning in higher education: a systematic review. *J. Technol. Sci. Educ.* 13:480. doi: 10.3926/jotse.1740
- Lewis, N. D. (2013). 100 statistical tests in R: What to choose, how to easily calculate, with over 300 illustrations and examples. Romeo, MI, USA: Heather Hills Press.
- Liu, T., Zhang, Z., and Gao, X. (Andy) (2023). Pedagogical Design in Technology-Enhanced Language Education Research: a scoping review. *Sustain. For.* 15:6069. doi: 10.3390/su15076069
- McDonald, R. P. (2009). Test theory: A unified treatment. New Jersey: Taylor & Francis e-library.
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika* 58, 525–543. doi: 10.1007/BF02294825/METRICKS
- Montaño-González, J., and Cancino, M. (2020). Exploring the relationship between language learning strategies and self-efficacy of Chilean university EFL students. *MEXTESOL Journal* 44. Available online at: <https://eric.ed.gov/?id=EJ1320831> (Accessed December 16, 2024).
- Montero-SaizAja, A. (2021). Gender-based differences in EFL learners' language learning strategies and productive vocabulary. *Theory Pract. Sec. Lang. Acquisit.* 7, 83–107. doi: 10.31261/TAPSLA.8594
- Oxford, R. L. (1986). Development and psychometric testing of the strategy inventory for language learning (SILL). Alexandria, VA: Research Institute for the Behavioral and Social Science.
- Oxford, R. L. (1990). Language learning strategies: What every teacher should know. Boston, MA: Heinle & Heinle Publishers.
- Oxford, R. L., and Burry-Stock, J. A. (1995). Assessing the use of language learning strategies worldwide with the ESL/EFL version of the strategy inventory for language learning (SILL). *System* 23, 1–23. doi: 10.1016/0346-251X(94)00047-A
- Peng, C. (2021). The academic motivation and engagement of students in English as a foreign language classes: does teacher praise matter? *Front. Psychol.* 12:778174. doi: 10.3389/fpsyg.2021.778174
- Putnick, D. L., and Bornstein, M. H. (2016). Measurement invariance conventions and reporting: the state of the art and future directions for psychological research. *Dev. Rev.* 41, 71–90. doi: 10.1016/J.DR.2016.06.004
- R Core Team (2023). R: a language and environment for statistical computing. Available online at: <https://cran.r-project.org> (Accessed May 4, 2023).
- Ranjan, R., and Philominraj, A. (2020). Language learning strategies, motivation and gender in foreign language context. *Univ. J. Educ. Res.* 8, 591–604. doi: 10.13189/ujer.2020.080231
- Revelle, W. (2022). Psych: procedures for psychological, psychometric, and personality research. Available online at: <https://cran.r-project.org/package=psych> (Accessed September 17, 2022).
- Rust, J., Kosinski, M., and Stillwell, D. (2021). Modern psychometrics. New York, NY, England: Routledge.
- Schumacker, R. E., and Lomax, R. G. (2016). A Beginner's guide to structural equation modeling. 4th Edn. New York: Taylor & Francis.
- Soper, D. (2024). A-priori sample size calculator for structural equation model. Available online at: <https://danielsoper.com/statcalc/calculator.aspx?id=89> (Accessed November 22, 2022).
- Sri Andayani, E. (2022). The importance of learning and knowing english in higher education in Indonesia. *Res. Dev. J. Educ.* 8, 372–379. doi: 10.30998/rdje.v8i1.13315
- Sumarni, S., and Rachmawaty, N. (2019). Gender differences in language learning strategies. *Ethical Lingua* 6, 13–22. doi: 10.30605/ethicallingua.v6i1.1169
- Tai, K. W. H., and Zhao, Y. V. (2024). Success factors for English as a second language university students' attainment in academic English language proficiency: exploring the roles of secondary school medium-of-instruction, motivation and language learning strategies. *Appl. Linguist. Rev.* 15, 611–641. doi: 10.1515/APPLIREV-2022-0049/
- Vandenberg, R. J., and Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: suggestions, practices, and recommendations for organizational research. *Organ. Res. Methods* 3, 4–70. doi: 10.1177/109442810031002
- Vieno, K., Rogers, K. A., and Campbell, N. (2022). Broadening the definition of 'research skills' to enhance students' competence across undergraduate and Master's programs. *Educ. Sci.* 12:642. doi: 10.3390/educsci12100642
- Wang, Y., Yasmin, F., and Akbar, A. (2023). Impact of the internet on English language learning among university students: mediating role of academic self-efficacy. *Front. Psychol.* 14:1184185. doi: 10.3389/fpsyg.2023.1184185
- Wei, L. (2023). Artificial intelligence in language instruction: impact on English learning achievement, L2 motivation, and self-regulated learning. *Front. Psychol.* 14:1261955. doi: 10.3389/fpsyg.2023.1261955
- Wen, Q., and Johnson, R. K. (1997). L2 learner variables and English achievement: a study of tertiary-level English majors in China. *Appl. Linguist.* 18, 27, 27–48. doi: 10.1093/applin/18.1.27
- Wilczewski, M., and Alon, I. (2023). Language and communication in international students' adaptation: a bibliometric and content analysis review. *High. Educ.* 85, 1235–1256. doi: 10.1007/s10734-022-00888-8
- Wu, S. C., and Chang, S.-M. (2018). Reclassifying the SILL: validation using exploratory and confirmatory factor analysis: An International Peer-reviewed Journal. Available online at: www.iiste.org (Accessed April 28, 2025).
- Yang, X., Zeng, L., and Xu, Z. (2021). An investigation of the language learning strategies used by Brazilian students learning Chinese as a foreign language. *Heliyon* 7:e07476. doi: 10.1016/j.heliyon.2021.e07476
- Zhai, X., and Razali, A. B. (2023). Triple method approach to development of a genre-based approach to teaching ESL/EFL writing: a systematic literature review by bibliometric, content, and Scientometric analyses. *SAGE Open* 13:255. doi: 10.1177/21582440221147255
- Zou, B., and Lertlit, S. (2022). Oxford's strategy inventory for language learning: English learning of Chinese students in Thai university. *LEARN J.* 15, 705–723.