Revista Educación, 2026, 50(2), julio-diciembre

ISSN: 0379-7082 / e-ISSN 2215-2644

Artículo científico de investigación

DOI: http://doi.org/10.15517/revedu.v50i2.686

Development, Validity, and Reliability of the Four Language Skills Self-Assessment Scale (FLSSAS) for Middle School Students

Desarrollo, Validez y Fiabilidad de la Escala de Autoevaluación de las Cuatro Habilidades Lingüísticas (FLSSAS) para Estudiantes de Secundaria

Arzu Gülbahçe
Atatürk University
Erzurum, Türkiye
arzugulbahce@atauni.edu.tr
https://orcid.org/0000-0003-4762-3603

Ayşen Kovan
Independent Researcher
Eskişehir, Türkiye
ayseen26@hotmail.com (Correspondencia)
https://orcid.org/0000-0002-4702-2675

Recepción: 02 de setiembre de 2025
Aceptado: 27 de abril de 2026

¿Cómo citar este artículo?

Gülbahçe, A., & Kovan, A. (2026). Development, Validity, and Reliability of the Four Language Skills Self-Assessment Scale (FLSSAS) for Middle School Students. *Revista Educación, 50*(2). http://doi.org/10.15517/revedu.v50i2.#ART

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

ABSTRACT

As English continues to function as a global and scientific lingua franca, supporting language learning through the development of both cognitive competence and affective confidence remains essential. During middle school, students begin to develop stable self-perceptions that significantly influence their motivation and engagement in language learning. However, no validated instrument currently exists to assess learners' competence and confidence across the four core English language skills—listening, speaking, reading, and writing—in English as a Foreign Language (EFL) contexts. This study aimed to develop and validate the Four Language Skills Self-Assessment Scale (FLSSAS), a self-report instrument designed to measure middle school students' competence and confidence in these four domains. A cross-sectional survey was conducted with 432 students aged 11 to 14 in Türkiye. An initial pool of 40 items was developed based on Bandura's Self-Efficacy Theory and Marsh's Self-Concept Theory. The instrument was subjected to exploratory and confirmatory factor analyses, and its internal consistency, composite reliability, average variance extracted and known-groups validity were evaluated. The results supported a two-factor structure (competence and confidence) within each language skill, demonstrating strong reliability and theoretical coherence. In addition, the FLSSAS exhibited developmental sensitivity across grade levels and robust construct validity. These findings indicate that the FLSSAS is a psychometrically sound instrument for assessing students' self-perceptions in English language learning and support a multidimensional framework applicable to both educational research and classroom assessment.

KEYWORDS: Competence, Confidence, Self-Efficacy, Four Language Skills, Scale Development.

RESUMEN

A medida que el inglés continúa sirviendo como una lengua franca tanto global como científica, apoyar el aprendizaje de idiomas mediante el fortalecimiento de la competencia cognitiva y la confianza afectiva es esencial. En la escuela secundaria, los estudiantes comienzan a formar autopercepciones estables que influyen significativamente en su motivación y participación en el aprendizaje de idiomas. Sin embargo, actualmente no existe ningún instrumento validado que evalúe la competencia y la confianza de los aprendices en las cuatro destrezas del idioma inglés (comprensión auditiva, expresión oral, comprensión lectora y expresión escrita) dentro de contextos de inglés como lengua extranjera (EFL, por sus siglas en inglés). Este estudio tuvo como objetivo desarrollar y validar la Escala de Autoevaluación de las Cuatro Habilidades Lingüísticas (FLSSAS, por sus siglas en inglés), un instrumento de autoinforme diseñado para medir la competencia y la confianza de los estudiantes de escuela secundaria en estos cuatro dominios. Se realizó un estudio transversal con 432 estudiantes de entre 11 y 14 años en Türkiye. Se generó un conjunto inicial de 40 ítems basados en la Teoría de la Autoeficacia de Bandura y la Teoría del Autoconcepto de Marsh. El instrumento fue sometido a aná-

lisis factorial exploratorio y confirmatorio, y se evaluó su consistencia interna, fiabilidad compuesta, varianza media extraída y validez de grupos conocidos. Los resultados respaldaron una estructura de dos factores (competencia y confianza) dentro de cada habilidad lingüística, demostrando una fuerte fiabilidad y coherencia teórica. Asimismo, la FLSSAS mostró sensibilidad al desarrollo a través de los distintos niveles escolares y una validez de constructo robusta. Estos hallazgos indican que la FLSSAS es un instrumento psicométricamente sólido para evaluar las autopercepciones de los estudiantes en el aprendizaje del idioma inglés y proponen un marco multidimensional aplicable tanto a la investigación educativa como a la evaluación en el aula

PALABRAS CLAVE: Competencia, Confianza, Autoeficacia, Cuatro habilidades lingüísticas, Desarrollo de escalas.

INTRODUCTION

In an era defined by global connectivity, the ability to communicate fluently in English, the world's de facto scientific and academic lingua franca, has become not merely advantageous but essential (Richards, 2015). English proficiency is now widely recognized as a foundational component of global literacy, shaping access to knowledge, professional opportunity, and intercultural dialogue. Accordingly, countries around the world, including Türkiye, have increasingly prioritized English language instruction across all levels of education, aiming to foster not only grammatical accuracy but communicative competence required for success in multilingual and multicultural contexts (Rose et al., 2021).

This growing emphasis raises important developmental questions particularly since middle school represents a pivotal stage in both cognitive and affective development. As learners transition from childhood to adolescence, they enter a psychologically formative period marked by deepening self-awareness, expanding abstract reasoning, and increasing motivation for autonomy and social identity (Erikson, 1968; Piaget, 1972). At this developmental threshold, students begin to form enduring beliefs about their academic abilities including those related to foreign language learning (Zimmerman, 2000). These beliefs act as both predictors and outcomes of educational engagement and are particularly influential in language learning, where success depends not only on mastering discrete skills but also on sustaining motivation, persistence, and confidence (Pavelescu & Petric, 2018; Wang & Bai, 2017).

In language instruction, these psychological processes unfold across four essential skill domains (listening, speaking, reading, and writing), which collectively form the foundation of communicative competence. While often taught concurrently, each skill imposes unique cognitive and emotional demands. Listening and reading emphasize receptive processing of incoming information, whereas speaking and writing require active construction and expression of meaning. How learners perceive

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

their abilities in these areas significantly influences their willingness to engage, persevere through challenges, and ultimately achieve language learning goals (Dörnyei & Ushioda, 2011).

Recent research has increasingly focused on the role of learner self-perception in these skill areas, particularly through the lenses of self-assessment and self-efficacy. Studies confirm that student self-perceptions, particularly as expressed through self-assessment and self-efficacy beliefs, significantly influence motivation, strategy use, and performance (Demirbulak et al., 2022; Mustafayeva, 2021; Özdemir & Karafil, 2023). Moreover, emerging research emphasizes the importance of distinguishing between cognitive and affective self-appraisals in second language acquisition, revealing that these dimensions contribute uniquely to learner development (Jamrus & Razali, 2019; Su, 2021; Taguchi & Roever, 2020; Zadorozhnyy & Lee, 2025).

To build a more detailed understanding of self-perception in language learning, it is essential to distinguish between two foundational psychological constructs: competence and confidence. Competence, grounded in academic self-concept theory (Marsh, 1990; Pajares & Schunk, 2002), refers to learners' cognitive evaluations of their ability to perform specific tasks. Confidence, based on Bandura's (1997) self-efficacy theory, reflects an individual's belief about their ability to succeed in a specific context. While closely related, these constructs are functionally distinct: for example, a student might believe they are capable of writing a paragraph in English (competence) yet still lack the confidence to attempt it. Together, these self-appraisals influence not only academic performance but also emotional regulation, goal-setting behavior, and persistence (Bong & Skaalvik, 2003; Papi & Teimouri, 2014). Recent studies conducted in English as a Foreign Language (EFL) contexts, including in Türkiye, have begun to operationalize this theoretical distinction in classroom-based research, showing, for instance, how learner self-efficacy mediates language use and willingness to communicate in informal or digital learning environments (Zadorozhnyy & Lee, 2025), and how students' confidence levels predict their engagement with oral communication strategies (Kansızoğlu, 2023; Su, 2021). These applied findings reinforce the need for assessment tools that not only reflect theoretical precision but also respond to classroom realities.

Despite the theoretical clarity of these concepts, assessment practices in EFL settings have not kept pace. Most language assessment tools do not differentiate between competence and confidence, particularly in EFL contexts. In Türkiye and similar contexts, assessment often remains teacher-centered, exam-driven, and grammar-focused (McKay, 2006), providing little insight into students' own perceptions of language development. Existing self-assessment instruments, where available, are typically narrow in scope and lack both multidimensional structure and psychometric rigor (Wang & Bai, 2017). This is especially problematic in middle school, when learners' academic self-beliefs are still forming and when accurate self-assessment can meaningfully shape motivation and achievement.

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

This study addresses that gap by presenting the development and validation of a new scale. Specifically, it introduces a theoretically grounded and psychometrically robust self-assessment instrument designed to capture middle school students' competence and confidence across the four essential domains of English: listening, speaking, reading, and writing. Grounded in educational psychology and second language pedagogy and aligned with the Common European Framework of Reference for Languages (CEFR) and the Turkish Ministry of National Education (MEB) curriculum, this instrument is intended to support both educational practice and empirical research. By integrating cognitive and affective dimensions of learner self-perception, the scale presents a more holistic and actionable framework for evaluating students' development in EFL contexts.

To guide the reader through this process, the remainder of the article is structured as follows. The methods section outlines the study design, participants, instrument development, and analytical procedures. The results section presents findings from exploratory and confirmatory factor analyses, as well as reliability, correlational, and grade-level comparisons. The discussion interprets these findings in light of relevant theories and research, addresses study limitations, and suggests directions for future inquiry. The article concludes by summarizing the contributions of the scale and its impact on language education and learner assessment.

METHODS

Design

This study used a quantitative, cross-sectional survey design to develop a self-report scale assessing middle school students' competence and confidence across four core English language skills. This design allowed for a snapshot of student perceptions across a developmentally diverse sample and is widely used in psychometric research (DeVellis, 2017). The study followed established scale development procedures, including: item generation, expert review and content validation, cognitive interviews with students, and large-scale administration for factor analysis and reliability testing (DeVellis, 2017; Worthington & Whittaker, 2006). The two main objectives were: (i) to create a valid, reliable item pool that captures both competence and skill-specific confidence, and (ii) to statistically examine the internal structure of the proposed scale using exploratory and confirmatory factor analyses. A self-assessment approach was selected to capture students' subjective evaluations of their abilities and emotions, in line with self-efficacy and academic self-concept theories, which emphasize the role of personal belief in learning engagement and motivation. A quantitative, cross-sectional design was preferred over qualitative or mixed-methods approaches due to its ability to systematically examine latent constructs across a large, developmentally varied sample. It offers strong psychometric precision essential for validating factor structure, reliability, and construct relationships while also enhancing replicability and generalizability, making the resulting scale both robust and widely applicable.

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

Procedure

Prior to data collection, ethical approval and all necessary institutional permissions were granted by the Social and Human Sciences Ethics Committee at Atatürk University (Approval No: E-56785782-050.02.04-2500142538, 05/18). Written informed consent was obtained from parents or guardians, and assent was secured from students. Participation was voluntary and confidential, with anonymity maintained in line with ethical standards for research involving minors.

The finalized scale was administered during regular school hours in classroom settings. Trained English teachers or research assistants gave brief instructions in Turkish, explaining that there were no right or wrong answers and encouraging honest responses. The paper-based survey took approximately 15–20 minutes to complete. Students who finished early were asked to remain seated quietly. Researchers or teachers monitored the process but did not explain or interpret item content. Completed scales were collected, anonymized, and entered into a secure database. Cases with over 10% missing data or uniform response patterns (e.g., same answer for every item) were excluded. Data collection was completed within two weeks to ensure classroom consistency.

Participants

The sample included 432 students (222 girls, 51.4%; 210 boys, 48.6%) from grades 5 through 8 in public middle schools located in a metropolitan district of Erzurum, Türkiye. This district was selected not only for accessibility and administrative cooperation, but also because it represents a linguistically and socioeconomically diverse population that reflects broader patterns in Turkish public education, including under-resourced school environments and regional disparities in foreign language exposure. These characteristics make it an important setting for understanding how middle school students evaluate their English language competence and confidence in less privileged educational contexts. Ages ranged from 11 to 14 years (M = 12.5, SD = 1.1). All students were native Turkish speakers, enrolled in schools following the national English language curriculum set by the MEB, which introduces English as a compulsory subject in Grade 4. By middle school, students typically receive 3–4 hours of English instruction per week, taught by licensed instructors.

Schools were selected using convenience sampling, based on accessibility and administrative approval. While this limits generalizability, it was appropriate for the exploratory nature of the study. The schools represented a socioeconomically diverse population, with most students coming from lower- to middle-income families.

Instrument Development

The Four Language Skills Self-Assessment Scale (FLSSAS) was designed to measure competence and confidence in each of the four language skill domains: listening, speaking, reading, and writing. Development was guided by the CEFR, the skill-based objectives of the MEB English curriculum,

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

and psychological theories on self-efficacy and academic self-concept (Bandura, 1997; Pajares & Schunk, 2002; Zimmerman, 2000). A comprehensive review of existing self-assessment tools and CEFR-based 'can-do' descriptors revealed no instrument that separately measured competence and confidence across all four skills at the middle school level. Thus, all items were newly constructed based on curriculum goals and age-appropriate language functions.

An initial pool of 40 items was created, with 10 items dedicated to each language skill domain. Each domain was deliberately designed to include two sub-dimensions: competence, reflecting students' ability (e.g., I can understand short English conversations), and confidence, reflecting self-efficacy and affective orientation toward the skill (e.g., I feel confident when listening to English). During item writing, each sub-dimension was operationalized separately to avoid conceptual overlap. Competence items focused on observable tasks linked to curriculum benchmarks (e.g., comprehension of instructions, participation in class activities), while confidence items targeted emotional and motivational responses (e.g., willingness to speak, fear of making mistakes). At least two sample items were drafted for each construct-skill intersection (e.g., speaking-confidence, reading-competence), and selections were made based on clarity and content coverage. Though some domains had unequal numbers of competence and confidence items (e.g., six vs. four), this imbalance reflected item quality after expert review, not theoretical prioritization. The goal was comprehensive coverage of both cognitive (skill-based) and affective (confidence-based) dimensions within each skill.

To ensure content validity, five experts (two English teachers, one applied linguist, one curriculum specialist, one educational psychologist) reviewed the items using a 5-point rubric assessing relevance, age appropriateness, and clarity. Items scoring low or flagged in comments were revised or removed. One item was deleted due to semantic overlap; six were revised for clarity (e.g., I *can answer English questions in class* → *I can answer English questions without much help*).

Following this, cognitive interviews were conducted with eight middle school students. Students were asked to think aloud as they answered selected items. Feedback revealed confusion on a few items (e.g., distinguishing *understanding English stories* vs. *finding the main idea*), prompting minor revisions. Overall, items were rated as clear and relatable by participants. All finalized items used a 5-point Likert-type scale (1 = Strongly disagree, 5 = Strongly agree). Higher scores reflected stronger competence or confidence, depending on the item. The full scale and all items are presented in English in Appendix A and in Turkish in Appendix B.

Data Analysis

Data were analyzed using IBM SPSS (v.23) and AMOS (v.23), following established guidelines for scale development in educational psychology (DeVellis, 2017; Worthington & Whittaker, 2006). Initial screening involved checking for missing data, extreme values, and response patterns. Cases

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

with more than 10% missing data or non-differentiated responses were excluded. Normality was assessed via skewness, kurtosis, and histograms. Kaiser-Meyer-Olkin (KMO) test and Bartlett's Test of Sphericity confirmed data suitability for factor analysis.

The total sample (N = 432) was randomly split into two equal subsamples (n = 216 each). Exploratory Factor Analysis (EFA) was conducted on the first subsample using principal axis factoring with promax rotation, run separately for each language skill. Items were grouped into competence and confidence subdimensions, guided by theory. Factors were extracted based on eigenvalues >1, scree plot inflection, and theoretical alignment. Items with loadings < .40 or cross-loadings were revised or removed.

Confirmatory Factor Analysis (CFA) was performed on the second subsample using maximum likelihood estimation. Each skill domain was modeled as a two-factor structure. Model fit was assessed using standard indices: $\chi^2/df < 3$, CFI and TLI ≥ .90, RMSEA ≤ .08, SRMR ≤ .08. Modifications were based on theoretical justification and modification indices. Full-scale CFA was not conducted due to the domain-specific structure of the instrument. Reliability was assessed using Cronbach's α, with .70 as the acceptability threshold. For smaller subscales, item-total and average inter-item correlations were reviewed. Pearson correlations were calculated across subscales to assess convergent and discriminant validity. To examine developmental trends, one-way ANOVA with Bonferroni post-hoc tests was conducted across grade levels. All statistical tests used a significance level of $p < .05$.

RESULTS

Exploratory Factor Analysis (EFA)

Separate EFAs were conducted for each language skill domain based on the theoretical structure comprising two sub-dimensions: competence and confidence. Principal axis factoring with promax rotation was used, assuming correlation between sub-dimensions.

Listening domain

Sampling adequacy was confirmed (KMO = .87), and Bartlett's test was significant (χ^2 = 945.28, df = 45, $p < .001$). A two-factor structure emerged, explaining 61.3% of total variance (competence = 42.8%, confidence = 18.5%). Factor loadings ranged from .58-.81 for competence and .74 - .79 for confidence, with no cross-loadings above .30 (see Table 1).

Table 1.

EFA Loadings – Listening Domain

Items	Competence Factor	Confidence Factor
I1. I can understand simple instructions in English.	.78	
I2. I can understand my teacher's English conversations.	.81	
I3. I can understand my classmates' English speech.	.76	

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

I4. I recognize the words I hear in English.	.73	
I5. I can understand short English conversations.	.77	
I6. I can understand English questions.	.70	
I7. I feel confident when listening to English.		.79
I8. I listen carefully when someone speaks English.		.74
I9. I can understand English in videos.	.72	
I10. I can guess the meaning of English words I don't know.	.58	

Source: Own elaboration, 2025.

Speaking domain

KMO = .85, and Bartlett's test was significant ($\chi^2 = 872.14$, df = 45, $p < .001$). Two factors explained 63.9% of variance (competence = 44.7%, confidence = 19.2%). Loadings ranged from .70-.81 for competence and .73 - .79 for confidence, with strong separation between sub-dimensions (Table 2).

Table 2.

EFA Loadings – Speaking Domain

Items	Competence Factor	Confidence Factor
I1. I can speak English fluently.	.76	
I2. I try to use new English words while speaking.	.81	
I3. I speak English often.	.74	
I4. I can speak English with simple sentences.	.70	
I5. I can answer English questions without much help.	.77	
I6. I can explain things about my daily life in English.	.72	
I7. I feel confident when speaking English in class.		.79
I8. I enjoy speaking English.		.75
I9. I try to participate in English conversations.		.78
I10. I am not afraid of making mistakes when speaking English.		.73

Source: Own elaboration, 2025.

Reading domain

KMO = .89, and Bartlett's test was significant ($\chi^2 = 1023.17$, df = 45, $p < .001$). The two factors explained 64.1% of variance (competence = 43.9%, confidence = 20.2%) competence item loadings ranged from .70-.79, and confidence items from .76-.78 (Table 3).

Table 3.

EFA Loadings – Reading Domain

Items	Competence Factor	Confidence Factor
I1. I can understand English texts.	.79	

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

I2. I can read and understand English instructions in activities.	.77	
I3. I recognize many words when reading English.	.76	
I4. I can find the main idea in a short paragraph.	.73	
I5. I understand what is happening when I read English stories.	.75	
I6. I can guess the meaning of words I don't know.	.74	
I7. I can read English quickly.	.70	
I8. I learn new English words while reading.	.72	
I9. I feel confident when reading short English texts.		.78
I10. I enjoy reading English.		.76

Source: Own elaboration, 2025.

Writing domain

KMO = .86, and Bartlett's test was significant ($\chi^2 = 981.65$, df = 45, $p < .001$). A two-factor solution explained 62.7% of variance (competence = 42.6%, confidence = 20.1%). Factor loadings were .70-.78 for competence and .77-.79 for confidence, confirming clear dimensional separation (Table 4).

Table 4.

EFA Loadings – Writing Domain

Items	Competence Factor	Confidence Factor
I1. I can write simple English sentences.	.78	
I2. I can use many English words when writing.	.75	
I3. I can express myself in short English sentences.	.76	
I4. I make few mistakes when writing in English.	.72	
I5. I can write about myself in English.	.73	
I6. I can do short writing homework assignments without much help.	.74	
I7. I can easily list what to write.	.71	
I8. I pay attention to the correct spelling and grammar of words when writing.	.70	
I9. I am confident when writing in English.		.79
I10. I enjoy writing in English.		.77

Source: Own elaboration, 2025.

Confirmatory Factor Analysis (CFA)

To confirm the factor structure identified through EFA, CFA was conducted for each domain to validate the two-factor model. Maximum likelihood estimation (MLE) was used. All CFA models

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

demonstrated acceptable to excellent fit, based on criteria of $\chi^2/df < 3$, CFI and TLI ≥ .90, RMSEA and SRMR ≤ .08. Model fit indices for each domain are presented in Table 5.

Table 5.

Model Fit Indices for CFA of Skill Domains

Domain	χ^2/df	CFI	TLI	RMSEA	SRMR
Listening	2.15	.95	.94	.06	.04
Speaking	2.41	.94	.93	.06	.05
Reading	2.10	.96	.95	.05	.04
Writing	2.33	.94	.93	.06	.05

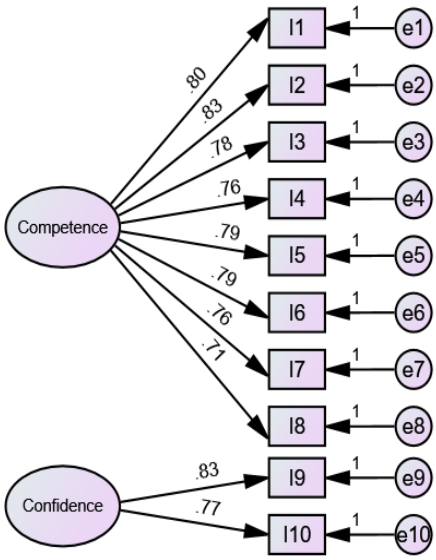
Source: Own elaboration, 2025.

Factor Loadings

All standardized CFA loadings were statistically significant ($p < .001$): in the listening domain, loadings for competence items ranged from .71-.83, while confidence items loaded between .77-.83, confirming a strong and well-defined two-factor structure (cf. Figure 1). In the speaking domain, competence item loadings ranged from .73 - .83, and confidence items loaded between .72-.81, again demonstrating a consistent alignment between observed indicators and their theorized sub-dimensions (cf. Figure 2).

Figure 1.

CFA Loadings in Listening

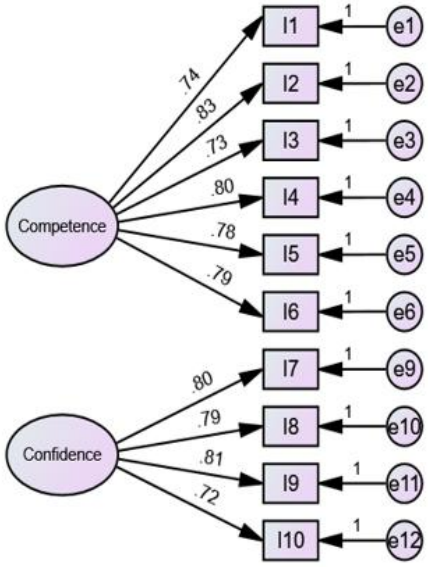


Source: Own elaboration, 2025.

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

Figure 2.

CFA Loadings in Speaking

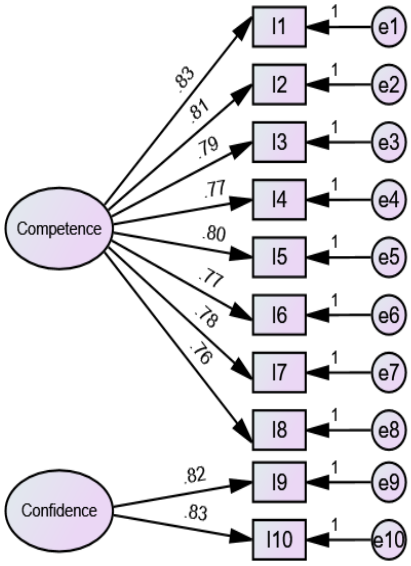


Source: Own elaboration, 2025.

For the reading domain, competence loadings ranged from .76-.83, while confidence items showed high loadings of .82-.83, supporting the validity of the hypothesized structure (cf. Figure 3). Finally, in the writing domain, competence items yielded loadings between .75-.85, and confidence items loaded at .81-.87, indicating excellent item-factor relationships for both sub-dimensions (cf. Figure 4).

Figure 3.

CFA Loadings in Reading

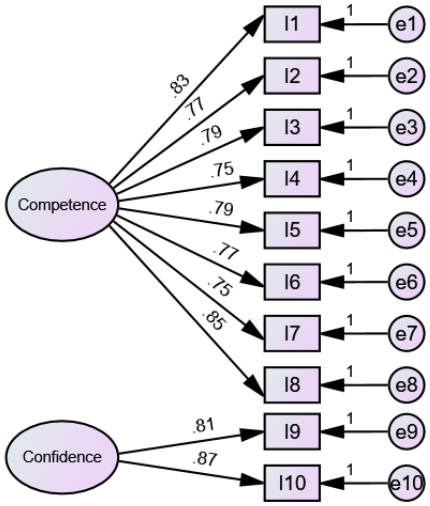


Source: Own elaboration, 2025.

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

Figure 4.

CFA Loadings in Writing



Source: Own elaboration, 2025.

These findings collectively demonstrate the robustness of the proposed two-factor model (competence and confidence) across each of the four language skill domains. The CFA results further confirm that the observed items reliably represent their theorized subdimensions, thereby providing strong evidence for the structural validity of the FLSSAS.

Reliability, Validity, and Descriptive Statistics

To evaluate the internal consistency and psychometric robustness of the FLSSAS, a comprehensive set of reliability and descriptive indices was calculated for each of the eight subscales representing competence and confidence dimensions within the four core skills. These results are shown in Table 6.

Table 6.

Reliability and Descriptive Statistics for Subscales

	1.	2.	3.	4.	5.	6.	7.	8.
1.Lis-Competence	—							
2.Lis-Confidence	.68	—						
3.Spe-Competence	.72	.66	—					
4.Spe-Confidence	.65	.74	.76	—				
5.Rea-Competence	.75	.70	.74	.69	—			
6.Rea-Confidence	.69	.73	.68	.76	.77	—		
7.Wri-Competence	.77	.71	.79	.75	.80	.74	—	
8.Wri-Confidence	.70	.78	.73	.81	.72	.79	.76	—
α	.89	.84	.88	.86	.91	.87	.92	.88

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

ω	.90	.86	.89	.87	.92	.88	.93	.89
λ2	.88	.83	.87	.85	.90	.86	.91	.87
CR	.91	.85	.90	.87	.93	.88	.94	.89
AVE	.63	.74	.68	.72	.70	.76	.71	.79
M	4.12	3.88	4.07	3.91	4.15	4.02	4.18	3.95
SD	.58	.65	.60	.63	.56	.59	.55	.63

Note. α: Cronbach's alpha, ω: McDonald's Omega, λ2: Guttman's Lambda-2, CR: Composite R liability, AVE: Average Variance Extracted values. All correlations are significant at $p < .001$

Source: Own elaboration, 2025.

Cronbach's α ranged from .84 to .92, exceeding the .70 threshold for all subscales (Nunnally & Bernstein, 1994). The highest α value was observed for Writing–Competence (α = .92), while Listening–Confidence yielded the lowest, though still satisfactory, value (α = .84). Additional reliability indices included McDonald's Omega (ω) values which ranged from .86 to .93, and Guttman's Lambda-2 (λ2) which ranged from .83 to .91. These confirm high internal consistency across both dimensions and skill domains.

Furthermore, Composite Reliability (CR) and Average Variance Extracted (AVE) values were derived from the CFA models to assess construct reliability and convergent validity. All CR values exceeded the .70 benchmark, ranging from .85 to .94, and all AVE values surpassed the .50 threshold, ranging from .63 to .79, thereby providing strong evidence of convergent validity (Hair et al., 2010). These results confirm that each latent construct is well represented by its respective indicators. Descriptive statistics also demonstrated mean values ranging from 3.88 to 4.18, suggesting that students generally reported high levels of both competence and confidence. The highest means were observed for Writing–Competence ($M = 4.18$, $SD = .55$) and Reading–Competence ($M = 4.15$, $SD = .56$). Standard deviation values, which ranged from .55 to .65, indicate an acceptable level of response variability without evidence of restricted range or skewed response patterns. Collectively, the reliability coefficients and descriptive findings provide compelling support for the psychometric adequacy, internal consistency, and convergent validity of the self-assessment scale across all skill domains and sub-dimensions.

Moreover, to further examine the construct validity of the scale, Pearson correlation coefficients were computed among the eight subscales representing Competence and Confidence within each of the four skill domains. As presented in Table 6, all correlations were positive and statistically significant ($p < .001$). The highest correlations were observed between Speaking–Confidence and Writing–Confidence ($r = .81$), and between Speaking–Competence and Writing–Competence ($r = .79$), suggesting strong interrelations between productive skills. Likewise, significant correlations were found within domains (e.g., Listening–Competence and Listening–Confidence, $r = .68$), supporting

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

the coherence of each of the dual dimensions of the domain. Importantly, while subscales were moderately to strongly correlated, no correlation exceeded .85, supporting discriminant validity. These patterns reinforce the theoretical distinction between competence and confidence while showing they are interrelated.

Grade-Level Comparisons

To assess known-groups validity, one-way ANOVA tests were conducted across Grades 5-8. Means, standard deviations, and F-statistics for each subscale are reported in Table 7.

Table 7.

Grade-Level Comparisons

	Grade 5	Grade 6	Grade 7	Grade 8	F	*p*
Lis-Compet	3.89 ± 0.55	4.05 ± 0.56	4.21 ± 0.52	4.28 ± 0.50	9.81	< .001
Lis-Confi	3.72 ± 0.59	3.90 ± 0.60	4.08 ± 0.61	4.16 ± 0.58	6.42	< .001
Spe-Compet	3.85 ± 0.52	4.03 ± 0.58	4.22 ± 0.56	4.30 ± 0.54	10.55	< .001
Spe-Confi	3.76 ± 0.60	3.92 ± 0.61	4.06 ± 0.59	4.18 ± 0.60	7.38	< .001
Rea-Compet	3.91 ± 0.57	4.12 ± 0.55	4.30 ± 0.54	4.35 ± 0.53	11.22	< .001
Rea-Confi	3.83 ± 0.54	4.00 ± 0.57	4.18 ± 0.55	4.25 ± 0.56	8.16	< .001
Wri-Compet	4.00 ± 0.51	4.15 ± 0.53	4.33 ± 0.52	4.40 ± 0.50	12.60	< .001
Wri-Confi	3.88 ± 0.56	3.92 ± 0.58	4.11 ± 0.57	4.20 ± 0.54	9.04	< .001

Source: Own elaboration, 2025.

Results showed statistically significant differences across all eight subscales ($p < .001$), suggesting that grade level has a meaningful influence on student self-assessment. In all domains, mean scores increased progressively with grade level, consistent with expected developmental trends in language acquisition and exposure to English instruction. For example, Writing–Competence showed the largest difference, increasing from $M = 4.00$ ($SD = 0.51$) in Grade 5 to $M = 4.40$ ($SD = 0.50$) in Grade 8 ($F(3, 428) = 12.60$, $p < .001$). Similarly, Reading–Competence improved steadily across grades ($F(3, 428) = 11.22$, $p < .001$), reflecting cumulative growth in comprehension skills. Post-hoc analyses (Bonferroni-adjusted) confirmed that students with higher grades reported significantly greater competence and confidence than their counterparts with lower grades, in nearly all comparisons. These findings provide strong evidence that the scale is sensitive to expected developmental differences, thus supporting its construct validity in educational settings.

DISCUSSION

This study was underpinned by two foundational psychological theories that together offer a detailed understanding of learner self-perception in educational contexts. First, Bandura's Self-Efficacy Theory (1997) conceptualizes confidence as the belief in one's ability to perform tasks successfully within a given domain. This captures motivational and affective dimensions of learning and student willingness

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

and persistence. Second, the theory of competence, closely tied to academic self-concept (Marsh, 1990), emphasizes the cognitive judgments of learners with regard to their language learning ability.

In language learning, the difference between *what I can do* (competence) and h*ow confident I feel about doing it* (confidence) reflects two distinct but interrelated psychological processes (Pajares & Schunk, 2002). This conceptual distinction directly informed the development and structure of the present scale. Recent research also supports this distinction in EFL settings, showing how beliefs about self-efficacy influence classroom engagement and goal-setting among adolescent learners (Wang & Bai, 2017). Empirical findings affirmed the validity of modeling competence and confidence as two distinct yet correlated latent constructs across the four core language skills. As shown in Tables 1, 2, 3 y 4, the EFA in each skill domain yielded a clear two-factor solution, with consistently high item loadings and minimal cross-loading. Notably, the confidence items in each domain formed distinct factors with loadings exceeding .73, affirming the theoretical structure of the scale. The strongest loading in Listening–Confidence and in Speaking–Confidence points to the clear emotional positioning of students when they engage with receptive and productive tasks, respectively.

Both exploratory and confirmatory factor analyses consistently supported a two-factor structure, each sub-dimension showing strong internal consistency. The factor structures are further supported by Figures 1, 2, 3 y 4, which consistently display strong path coefficients, thereby visually reinforcing the conceptual split between competence and confidence. Table 5 further supports model validity, with all fit indices falling within acceptable thresholds, reinforcing the structural soundness of the proposed two-factor model across all domains. The CFA model for writing showed the highest standardized loadings, indicating a particularly stable perception of writing-related self-efficacy. Students were able to distinguish between their capability to complete a task (competence) and their belief in success (confidence) (Pajares & Schunk, 2002; Talsma et al., 2018; Teng & Zhang, 2016). Importantly, this differentiation was robust across all language domains, suggesting a generalized self-perception framework that transcends skill type. This pattern aligns with recent case studies that demonstrate how adolescent learners form distinct emotional associations with different language skills, particularly in EFL classrooms (Pavelescu & Petric, 2018).

One of the most notable findings was the strong coherence among confidence items, which consistently loaded onto their own factor, independently from competence. The pattern visible in Table 6 confirms that internal consistency measures for confidence subscales were all above .83, indicating that affective appraisals were both reliable and systematically distinct from student cognitive evaluations. Notably, the Writing–Competence subscale showed the highest internal consistency across all indices, suggesting that students may feel particularly sure of their written output, possibly because writing allows for reflection, drafting, and teacher feedback. In contrast, Listening–Confidence yiel-

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

ded the lowest consistency values, which may reflect the fleeting and high-pressure structure of listening tasks, where learners have limited control over input pace or repetition. This supports Bandura's assertion that self-efficacy operates as a motivational force, separate from actual ability (Bandura, 1997; Zimmerman, 2000). The coexistence of high confidence with low competence, and vice versa, highlights a common but often overlooked pattern of divergence between emotional and cognitive self-appraisals (Eccles & Wigfield, 2002; Marsh, 1990; Pajares, 1996). Additionally, the competence subscales displayed developmental sensitivity. This trend, seen most clearly in Table 7, illustrates significant increases in competence and confidence across all skill areas from Grades 5 through 8, with the most pronounced growth observed in the writing and reading domains. Notably, Writing–Competence showed the largest increase from Grade 5 to Grade 8, likely reflecting the cumulative nature of writing instruction and students' growing exposure to structured practice over time. In contrast, Listening–Confidence showed the smallest gains, which may indicate persistent challenges in real-time language processing or a relatively lower instructional emphasis on listening strategies. These patterns support the study objective of identifying skill-specific developmental trends and suggest valuable insight for educators seeking to tailor support to students' evolving self-perceptions. Longitudinal comparisons showed steady improvement across grade levels, indicating that as students gain experience, their self-assessments become more refined and accurate, consistent with academic self-concept theory (Guay et al., 2003; Marsh, 1990; Pinxten et al., 2010). This reflects meaningful psychological growth during early adolescence (Burns et al., 2020; Eccles & Roeser, 2011). For example, classroom-based research in Türkiye has shown that consistent exposure to writing instruction significantly shapes student confidence and competence in composition tasks, especially in upper middle school grades (Kansızoğlu, 2023).

Although distinct, competence and confidence were positively correlated, affirming their conceptual interdependence. Pearson correlations among the eight subscales (Table 6) further confirm this: strong within-domain correlations (e.g., listening–competence and listening–confidence) validate the internal coherence of each skill area, while moderate cross-domain relationships highlight the broader interconnectedness of language self-perception across modalities. Confidence, while informed by competence, remains an affective construct with its own role in shaping behavior (Bong & Skaalvik, 2003; Pajares, 1996). This theoretical separation, while maintaining relational links, allows for a more detailed interpretation of learner self-perception.

In effect, the study provides strong evidence that assessing these constructs separately offers richer insights than undifferentiated self-assessment tools (Bong & Skaalvik, 2003; Marsh & Craven, 2006). The validated scale structure not only reinforces key theoretical positions but also offers a practical tool aligned with curriculum standards. Teachers can use the scale formatively to identi-

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

fy students who, for example, report high competence but low confidence, highlighting a need for affective support such as positive feedback or modeling. Similarly, learners with high confidence but low competence may benefit from targeted skill-building tasks to align their self-beliefs with actual performance. It enables educators to better identify discrepancies between belief and ability and design targeted interventions, particularly in the English as a Foreign Language (EFL) context. Because the instrument uses age-appropriate language and aligns with the MEB English curriculum, it is especially useful for middle school learners, an age when self-perceptions are still forming. This is consistent with current pedagogical perspectives that advocate for differentiated instruction based on learner beliefs and identity in Global Englishes contexts (Rose et al., 2021). Practically, the scale can be integrated at the beginning and end of academic terms to monitor changes in self-perception, guide instructional decisions, and tailor support to specific skills like listening or writing where learners often diverge in confidence and competence. Furthermore, the tool holds promise for researchers examining how the affective and cognitive components of language learning interact and evolve across time or instructional settings. Its domain-specific structure also enables comparisons between productive and receptive skills, offering insights into how student self-perceptions develop differently depending on the type of language task. It bridges theoretical and applied domains, offering a replicable model for both psychometric research and educational practice in second language acquisition.

Despite its strengths, this study is not without limitations. Initially, while the study provides valuable insights into middle school students in Eastern Türkiye, the use of a geographically limited, convenience-based sample may constrain the generalizability of the findings. Future studies should aim to validate the scale in varied national and international contexts, including multilingual or immersion-based educational settings, to assess the robustness and cultural adaptability of its constructs (Wang & Bai, 2017). This includes both high-resource and under-resourced school systems, where students' self-perceptions may differ based on access to instruction and materials. Moreover, the scale is based entirely on self-report, which, although appropriate given the focus on competence and confidence, introduces the possibility of response bias, particularly social desirability or limited introspective accuracy in younger respondents. Future research should also triangulate self-report data with additional sources such as teacher-rated rubrics, classroom observations, or performance-based assessments, thereby strengthening construct validity through multi-method evidence. Besides, while the factor structure was confirmed via CFA, longitudinal data were not collected, meaning that scale sensitivity to developmental change over time remains unexplored. Future research could incorporate repeated measures across academic years to assess how instructional practices, curricular shifts, or individual learner trajectories influence changes in confidence and competence over time. Finally, although the instrument was designed in alignment with the CEFR and the MEB curriculum, its

applicability to other language education frameworks or age groups has yet to be tested. Exploring its use in upper-secondary or adult learner populations, as well as in informal language learning environments, would further extend its utility and relevance. Ultimately, the scale offers a practical bridge between psychological theory and classroom application, allowing educators not only to measure learner perceptions but also to actively use these data to shape more responsive, differentiated English language instruction. These limitations highlight the need for continued psychometric and contextual refinement to further establish the robustness and versatility of the scale.

CONCLUSION

This study introduced and validated the FLSSAS, a theoretically grounded and psychometrically robust instrument designed to assess middle school student competence and confidence in English listening, speaking, reading, and writing. Drawing on the Self-Efficacy Theory by Bandura and the Self-Concept Theory by Marsh, the scale addresses a critical need in language education for tools that account for both the cognitive and affective dimensions of learner self-perception. Through a rigorous process of item development, expert validation, exploratory and confirmatory factor analyses, and multiple reliability indicators, this study established that competence and confidence are empirically distinct yet interrelated constructs. The scale not only reflects the structure of the Turkish national curriculum, but also captures developmental trends, as evidenced by meaningful grade-level differences across all subscales.

Beyond its methodological rigor, the study makes a significant conceptual contribution by proposing a multidimensional framework for self-assessment in second language learning. By disentangling what students believe they can do from how confident they feel about executing it, the scale provides a more detailed and actionable understanding of language learner psychology. This differentiation opens new avenues for classroom practice, assessment design, and research into learner development and motivation. As a contextually relevant, developmentally appropriate, and theoretically sound tool, the scale stands to benefit educators, curriculum designers, and researchers interested in fostering reflective, confident, and capable language learners. In doing so, it lays important groundwork for expanding how learner perceptions are measured, interpreted, and supported in foreign language education.

REFERENCES

Bandura, A. (1997). *Self-efficacy: The exercise of control*. W. H. Freeman.

Bong, M., & Skaalvik, E. M. (2003). Academic self-concept and self-efficacy: How different are they really? *Educational Psychology Review, 15*(1), 1-40. https://doi.org/10.1023/A:1021302408382

Burns, R. A., Crisp, D. A., & Burns, R. B. (2020). Re-examining the reciprocal effects model of self-concept, self-efficacy, and academic achievement in a comparison of the cross-

lagged panel and random-intercept cross-lagged panel frameworks. *British Journal of Educational Psychology, 90*(1), 77-91. https://doi.org/10.1111/bjep.12265

Demirbulak, D., Icden, L., & Yılmaz-Virlan, A. (2022). A case study on the effect of self-assessment on EFL learners' self-efficacy beliefs. *European Journal of English Language Studies, 2*(1), 35-46. https://doi.org/10.12973/ejels.2.1.35

DeVellis, R. F. (2017). *Scale development: Theory and applications* (4th ed.). Sage.

Dörnyei, Z., & Ushioda, E. (2011). *Teaching and researching motivation* (2nd ed.). Routledge.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence, 21*(1), 225-241. https://doi.org/10.1111/j.1532-7795.2010.00725.x

Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*, 109–132. https://doi.org/10.1146/annurev.psych.53.100901.135153

Erikson, E. H. (1968). *Identity: Youth and crisis.* Norton.

Guay, F., Marsh, H. W., & Boivin, M. (2003). Academic self-concept and academic achievement: Developmental perspectives on their causal ordering. *Journal of Educational Psychology, 95*(1), 124-136. https://doi.org/10.1037/0022-0663.95.1.124

Hair, J. F., Black, W. C., Balin, B. J., & Anderson, R. E. (2010). *Multivariate data analysis.* Maxwell Macmillan International Editions.

Jamrus, M. H. M., & Razali, A. B. (2019). Using self-assessment as a tool for English language learning. *English Language Teaching, 12*(11), 64-73. https://doi.org/10.5539/elt.v12n11p64

Kansızoğlu, H. B. (2023). Investigating Turkish teachers' views and practices on writing instruction in secondary schools: A mixed-methods study. *Journal of Pedagogical Research, 7*(3), 211-247. https://doi.org/10.33902/JPR.202319728

Marsh, H. W. (1990). A multidimensional, hierarchical model of self-concept: Theoretical and empirical justification. *Educational Psychology Review, 2*(2), 77–172. https://doi.org/10.1007/BF01322177

Marsh, H. W., & Craven, R. G. (2006). Reciprocal effects of self-concept and performance from a multidimensional perspective: Beyond seductive pleasure and unidimensional perspectives. *Perspectives on Psychological Science, 1*(2), 133-163. https://doi.org/10.1111/j.1745-6916.2006.00010.x

McKay, P. (2006). *Assessing young language learners*. Cambridge University Press.

Mustafayeva, N. (2021). The effect of self assessment on the EFL students' self efficacy concerning speaking skills. *International Journal of Social and Humanities Sciences Research (JSHSR), 8*(75), 2617-2621. https://doi.org/10.26450/jshsr.2473

Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric theory* (3rd ed.). McGraw-Hill.

Özdemir, O., & Karafil, B. (2023). Exploring English self-efficacy beliefs of university students based on EFL students' and their instructors' opinions. *Gazi Üniversitesi Gazi Eğitim Fakültesi Dergisi, 43*(3), 2237-2271. https://doi.org/10.17152/gefad.1275454

Pajares, F. (1996). Self-efficacy beliefs in academic settings. *Review of Educational Research, 66*(4), 543–578. https://doi.org/10.3102/00346543066004543

Pajares, F., & Schunk, D. H. (2002). Self and self-belief in psychology and education: A historical perspective. In J. Aronson (Ed.), *Improving academic achievement: Impact of psychological factors on education* (pp. 3–21). Academic Press. https://doi.org/10.1016/B978-012064455-1/50004-X

Papi, M., & Teimouri, Y. (2014). Language learner motivational types: A cluster analysis study. *Language Learning: A Journal of Research in Language Studies, 64*(3), 493-525. https://doi.org/10.1111/lang.12065

Pavelescu, L. M., & Petrić, B. (2018). Love and enjoyment in context: Four case studies of adolescent EFL learners. *Studies in Second Language Learning and Teaching, 8*(1), 73-101. https://doi.org/10.14746/ssllt.2018.8.1.4

Piaget, J. (1972). *The psychology of the child.* Basic Books.

Pinxten, M., De Fraine, B., Van Damme, J., & D'Haenens, E. (2010). Causal ordering of academic self-concept and achievement: Effects of type of achievement measure. *British Journal of Educational Psychology, 80*(4), 689-709. https://doi.org/10.1348/000709910X493071

Richards, J. C. (2015). *Key issues in language teaching.* Cambridge University Press.

Rose, H., McKinley, J., & Galloway, N. (2021). Global Englishes and language teaching: A review of pedagogical research. *Language Teaching, 54*(2), 157-189. https://doi.org/10.1017/S0261444820000518

Su, Y. C. (2021). College students' oral communication strategy use, self-perceived English proficiency and confidence, and communication anxiety in Taiwan's EFL learning. *Educational Studies, 57*(6), 650-669. https://doi.org/10.1080/00131946.2021.1919677

Taguchi, N., & Roever, C. (2020). *Second language pragmatics*. Oxford University Press.

Talsma, K., Schüz, B., Schwarzer, R., & Norris, K. (2018). I believe, therefore I achieve (and vice versa): A meta-analytic cross-lagged panel analysis of self-efficacy and academic performance. *Learning and Individual Differences, 61,* 136–150. https://doi.org/10.1016/j.lindif.2017.11.015

Teng, L. S., & Zhang, L. J. (2016). A questionnaire-based validation of multidimensional constructs of self-regulated learning in college writing. *The Modern Language Journal, 100*(3), 674–701. https://doi.org/10.1111/modl.12339

Wang, C., & Bai, B. (2017). Validating the instruments to measure ESL/EFL learners' self-efficacy beliefs and self-regulated learning strategies. *TESOL Quarterly, 51*(4), 931-947. https://doi.org/10.1002/tesq.355

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

Worthington, R. L., & Whittaker, T. A. (2006). Scale development research: A content analysis and recommendations for best practices. *The Counseling Psychologist, 34*(6), 806-838. https://doi.org/10.1177/0011000006288127

Zadorozhnyy, A., & Lee, J. S. (2025). Informal digital learning of English and willingness to communicate in a second language: Self-efficacy beliefs as a mediator. *Computer Assisted Language Learning, 38*(4), 669-689. https://doi.org/10.1080/09588221.2023.2215279

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82–91. https://doi.org/10.1006/ceps.1999.1016

APPENDIXES

Appendix A.

Items of the Four Language Skills Self-Assessment Scale (FLSSAS)

Explanation: The FLSSAS was developed to evaluate middle school students' levels of competence and confidence in the four core areas of English (listening, speaking, reading, and writing). Each item is rated on a 5-point Likert scale: 1 = Strongly disagree, 2 = Disagree, 3 = Neutral, 4 = Agree, 5 = Strongly agree. Higher scores indicate a higher level of competence and/or confidence in the corresponding language skill. The sub-dimensions of competence and confidence are addressed in the EFA and CFA sections.

Domains	Items	1	2	3	4	5
Listening	I can understand simple instructions in English.					
	I can understand my teacher's English conversations.					
	I can understand my classmates' English speech.					
	I recognize the words I hear in English.					
	I can understand short English conversations.					
	I can understand English questions.					
	I feel confident when listening to English.					
	I listen carefully when someone speaks English.					
	I can understand English in videos.					
	I can guess the meaning of English words I don't know.					
Speaking	I can speak English fluently.					
	I try to use new English words while speaking.					
	I speak English often.					
	I can speak English with simple sentences.					
	I can answer English questions without much help.					
	I can explain things about my daily life in English.					
	I feel confident when speaking English in class.					
	I enjoy speaking English.					
	I try to participate in English conversations.					

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

	I am not afraid of making mistakes when speaking English.
Reading	I can understand English texts.
	I can read and understand English instructions in activities.
	I recognize many words when reading English.
	I can find the main idea in a short paragraph.
	I understand what is happening when I read English stories.
	I can guess the meaning of words I don't know.
	I can read English quickly.
	I learn new English words while reading.
	I feel confident when reading short English texts.
	I enjoy reading English.
Writing	I can write simple English sentences.
	I can use many English words when writing.
	I can express myself in short English sentences.
	I make few mistakes when writing in English.
	I can write about myself in English.
	I can do short writing homework assignments without much help.
	I can easily list what to write.
	I pay attention to the correct spelling and grammar of words when writing.
	I am confident when writing in English.
	I enjoy writing in English.

Appendix B.

Dört Dil Becerisi Öz-Değerlendirme Ölçeği (DDBÖDÖ) Maddeleri

Açıklama: DDBÖDÖ, ortaokul öğrencilerinin İngilizcenin dört temel alanındaki (dinleme, konuşma, okuma ve yazma) yeterlik (competence) ve özgüven (confidence) düzeylerini değerlendirmek amacıyla geliştirilmiştir. Her bir madde 5'li Likert ölçeğinde derecelendirilmiştir: 1 = Kesinlikle katılmıyorum, 2 = Katılmıyorum, 3 = Kararsızım, 4 = Katılıyorum, 5 = Kesinlikle katılıyorum. Daha yüksek puanlar, ilgili dil becerisinde daha yüksek düzeyde yeterlik ve/veya özgüveni göstermektedir. Yeterlik ve özgüven alt boyutları, AFA ve DFA bölümlerinde ele alınmıştır.

Alanlar	Maddeler	1	2	3	4	5
Dinleme	İngilizce verilen basit yönergeleri anlayabiliyorum.					
	Öğretmenimin İngilizce konuşmalarını anlayabiliyorum.					
	Sınıf arkadaşlarım İngilizce konuştuğunda ne dediklerini anlayabiliyorum.					
	İngilizce duyduğum kelimeleri tanıyorum.					

Revista Educación, 2026, 50(2), julio-diciembre ISSN: 0379-7082 / e-ISSN 2215-2644

	Kısa İngilizce konuşmaları anlayabiliyorum.
	İngilizce soruları anlayabiliyorum.
	İngilizce dinlerken kendime güveniyorum.
	Birisi İngilizce konuştuğunda dikkatli dinliyorum.
	İngilizce sesli ya da görüntülü içerikleri izlerken ne söylendiğini anlayabiliyorum.
	Bilmediğim İngilizce kelimelerin anlamını tahmin edebiliyorum.
Konuşma	İngilizceyi akıcı bir şekilde konuşabiliyorum.
	Konuşurken yeni İngilizce kelimeler kullanmaya çalışıyorum.
	İngilizceyi sık sık konuşuyorum.
	Basit cümlelerle İngilizce konuşabiliyorum.
	İngilizce sorulara çok yardım almadan cevap verebiliyorum.
	Günlük hayatımla ilgili şeyleri İngilizce anlatabiliyorum.
	Sınıfta İngilizce konuşurken kendime güveniyorum.
	İngilizce konuşmaktan keyif alıyorum.
	İngilizce konuşmalara katılmaya çalışıyorum.
	İngilizce konuşurken hata yapmaktan korkmuyorum.
Okuma	İngilizce metinleri anlayabiliyorum.
	Etkinliklerdeki İngilizce yönergeleri okuyup anlayabiliyorum.
	İngilizce okurken birçok kelimeyi tanıyorum.
	Kısa bir paragrafta ana fikri bulabiliyorum.
	İngilizce hikâyeleri okuyunca ne olduğunu anlıyorum.
	Bilmediğim kelimelerin anlamını tahmin edebiliyorum.
	İngilizceyi hızlı okuyabiliyorum.
	Okurken yeni İngilizce kelimeler öğreniyorum.
	Kısa İngilizce metinleri okurken kendime güveniyorum.
	İngilizce okumaktan keyif alıyorum.
Yazma	Basit İngilizce cümleler yazabiliyorum.
	Yazarken birçok İngilizce kelime kullanabiliyorum.
	Kısa İngilizce cümlelerle kendimi ifade edebiliyorum.
	İngilizce yazarken az hata yapıyorum.
	Kendimle ilgili şeyleri İngilizce yazabiliyorum.
	Kısa yazma ödevlerini çok fazla yardım almadan yapabiliyorum.
	Ne yazacağımı kolayca sıralayabiliyorum.
	Yazarken kelimelerin doğru yazımına ve dilbilgisine dikkat ediyorum.
	İngilizce yazarken kendime güveniyorum.
	İngilizce yazı yazmaktan keyif alıyorum.