Adv Mobile Learn Educ Res, 2023, 3(2): 787-800
DOI: 10.25082/AMLER.2023.02.006

RESEARCH ARTICLE

Development of a Scale to Measure Children's Educational Mobile Application Usage for Foreign Language

Derya Uygun[1*] **Mehmet Fırat**[1]

[1] Department of Distance Education, Faculty of Open Education, Anadolu University, Eskisehir, Turkey

Correspondence to: Derya Uygun, Department of Distance Education, Faculty of Open Education, Anadolu University, Eskisehir, Turkey;
Email: duygun@anadolu.edu.tr

Received: May 8, 2023;
Accepted: July 27, 2023;
Published: August 2, 2023.

Citation: Uygun, D., & Fırat, M. (2023). Development of a Scale to Measure Children's Educational Mobile Application Usage for Foreign Language. *Advances in Mobile Learning Educational Research*, 3(2), 787-800. https://doi.org/10.25082/AMLER.2023.02.006

Abstract: The widespread use of mobile applications in education has led to the need for reliable and valid measurement tools to assess their impact on language learning, particularly among young learners. This study investigated the developmental process and psychometric properties of the Children's Educational Mobile Application Usage for Foreign Languages Scale. The item pool was created by examining the studies obtained from the literature review. To evaluate the appearance and content validity of the scale, the expert assessment of the scale, which the researcher developed, was applied to 19 experts. Consequently, the experts' opinions were considered to update the Children's Educational Mobile Application Usage for Foreign Languages Scale. Pilot studies were conducted with 37 students in the 4th grade. Data were collected from 309 4th-grade students to conduct reliability and validity studies. Confirmatory Factor Analysis was used to determine the scale's construct validity. The Chi-square/degrees of freedom ratio is 2.54, and the Comparative Fit Index value is calculated as 0.954. The Confirmatory Factor Analysis confirmed the theoretical structure consisting of 16 items with three factors, and the compliance indices showed perfect compliance. Cronbach Alpha (α) coefficient was used to determine the scale's reliability. The Cronbach Alpha value is calculated as 0.945. As a result, the Children's Educational Mobile Application Usage for Foreign Languages Scale was valid and reliable. The study highlights the novelty of the scale, being one of the first psychometric scales in the literature explicitly targeting children's educational mobile application usage for foreign language learning. It can aid institutions in developing or updating mobile applications and investigating the relationship between children's educational mobile application use and other variables. While this study presents a valuable contribution to the field, it also acknowledges limitations, such as the sample size and reliance on personal statements for data collection. Nonetheless, Children's Educational Mobile Application Usage for Foreign Languages opens new avenues for research and has the potential to enhance language education by harnessing innovative technology-driven approaches.

Keywords: mobile learning, mobile assisted language learning, foreign language development of children, autonomy, factor analysis

1 Introduction

The opportunities offered by technology are among the complementary elements in foreign language learning and skill acquisition activities (Sutrisni et al., 2022). There has been a shift from Computer-Assisted Language Learning (CALL) to Mobile-Assisted Language Learning (MALL). Many applications designed for CALL have now been transferred to MALL (Alkan, 2011; Yang, 2013). MALL has started to take students beyond the confines of the classroom to the real world. Traditional teaching methods may not meet the expectations of the new generation of digital natives, as they are already familiar with the technology before starting preschool education (Papadakis et al., 2021; Sunar, Dahal & Pant, 2022). Although some parents may delay introducing technological devices to their children, many provide technology almost unlimitedly (Strataki, 2022). Thus, it is inevitable for the new generation, almost surrounded by technology, to incorporate it into their formal education (Papadakis et al., 2022).

The age at which foreign language learning begins can be considered among the crucial parameters in language acquisition. The fundamental reason for this is that individuals who start learning a foreign language at a young age are seen to be more successful than adults (Johnson & Newport, 1991; Nunan, 2003; Slavoff & Johnson, 1995). Therefore, starting foreign language education from a young age may positively impact children's language skills development and increase the use of educational mobile applications (Papadakis et al., 2022). However, children's language skills and learning processes can vary depending on age (Dockrell & Marshall, 2015). For example, younger children, who are five years old, may prefer educational

mobile applications that focus on developing basic language grammar and vocabulary. In comparison, older children, who are ten years old, may show interest in applications that enhance more complex language structures and communication skills. Therefore, selecting and using educational mobile applications suitable for children's ages can be crucial in supporting their language learning processes (Skaraki, 2021; Tsoukala, 2021).

The impact of mobile technologies on foreign language acquisition is increasing daily due to their convenience and accessibility from anywhere at any time (Munajah, Sumantri, & Yufiarti, 2022). Utilizing mobile applications on portable devices is one way that positively affects students' foreign language learning (Papadakis & Kalogiannakis, 2020; 2022). Mobile applications can support students individually or collaboratively in reading, listening, speaking, and writing in the target language. MALL can bridge in-class and out-of-class learning in formal and informal settings.

It is observed that there are many paid or free mobile applications that support the process of learning English as a foreign language. However, it is noteworthy that the studies about the effectiveness of these applications on children's foreign language development need to be revised (Booton et al., 2023). Because reliable and valid scales to evaluate the effectiveness of these apps on children's foreign language development are inadequate. To establish a link between theory and practice, the effectiveness of the applications should be investigated first and then disseminated (Torres, 2023).

In this context, there is a significant need for a scale to evaluate the effects of mobile applications supporting foreign language learning on 4th-grade students' development. This paper presents a scale, the Children's Educational Mobile Application Usage for Foreign Language, to determine the effectiveness of such mobile applications.

1.1 Review of scales used in determining effectiveness of mobile applications

When the literature is examined, it is observed that various scales are used to determine the effectiveness of mobile applications (Önal et al., 2022). One of them is the user satisfaction scale. These scales measure the satisfaction levels of users with the mobile application. They may ask users to rate their overall satisfaction with the app, its interface, ease of use, and perceived usefulness. Sözeri and Harmanşah (2018) developed a mobile application called Learning Training Set English (LETS English) to support English lessons. The researchers conducted a study that included learners' opinions about the application. As a result, it was reported that 28 students who used the application had a positive view of it. When researchers develop an application, it is suggested that they assess its effectiveness using valid and reliable measurement tools.

Despite the significance of measuring user satisfaction, only some scientific studies focused on scales' development, reliability, and validity. The lack of standardized and rigorously tested scales can lead to consistent and reliable results, hindering the accurate assessment of user satisfaction. Putra and Law (2023) further stressed the importance of explicitly developing and validating scales tailored for mobile applications. They argued that generic scales, commonly used in other contexts, may need to fully capture the nuances and intricacies of mobile app user experiences, potentially leading to misleading conclusions and recommendations.

The study of Fernandes and Pedroso (2017) highlights the need for more comprehensive research in this domain. They emphasized that the existing scales used to measure user satisfaction may need to be adequately adapted to the context of mobile applications, which often require unique assessment approaches due to their specific user interactions and features.

Another type of scale used in the effectiveness of mobile applications is the learning achievement scale. These scales measure the learning outcomes achieved by users after using the mobile application. These achievement tests can be administered before and after users interact with the mobile application to assess the app's effectiveness in improving their language skills and overall learning outcomes.

Researchers generally use self-developed achievement tests to determine the effectiveness of mobile applications. Wu (2015) developed a mobile application called Word Learning-CET6 that can be used in English vocabulary teaching for Chinese students and investigated the effectiveness of this application. The researcher developed the word test as a data collection tool. As a result of the research, it has been reported that the mobile application, Word Learning-CET6, is effective in teaching English vocabulary. It is essential to design the tests carefully to ensure their reliability and validity in measuring learning achievements accurately.

Achievement tests may not capture the full range of learning outcomes or the overall impact of the mobile app. They often focus on specific language skills or knowledge, potentially neglecting other essential aspects of learning (Sato, 2023). Moreover, achievement tests may

only be specific to the application used in the research. No valid and reliable success test covers all applications. While various scales are mentioned, there is a need for standardized scales exclusively developed for mobile applications to ensure consistency and comparability across different studies (Putra & Law, 2023).

Another type of scale used in the effectiveness of mobile applications is the Perceived Learning Scale. These scales assess learners' perceptions of the educational value and effectiveness of the mobile application. They can gauge whether users believe the app has contributed to their learning and language skills (Wang et al., 2023). These scales can assess various aspects of learning, including perceived improvements in language skills, knowledge acquisition, confidence, and motivation. They provide a more comprehensive evaluation of the app's effect on learners' overall learning experience (Wang et al., 2023). However, no valid and reliable perceived learning scale is developed only for mobile applications.

Children's Educational Mobile Application Usage for Foreign Language (CEMAUFL) will be a first in the literature. The scale can be used to determine the effect of mobile applications that support foreign language learning on the development of 4th-grade students in English lessons. The scale will make an essential contribution to the literature. The scale is vital for researchers in foreign languages interested in researching mobile applications. CEMAUFL aids researchers in understanding the impact of mobile apps on language learning for 4th-grade students.

1.2 Aim of the study

The article aims to develop and validate a scale called "Children's Educational Mobile Application Usage for Foreign Language" to assess the usage of mobile applications among 4th-grade students for foreign language learning. The study focuses on measuring the impact of mobile applications on children's English language skills and explores the effectiveness of these applications in enhancing language learning outcomes. The article aims to establish the psychometric properties, validity, and reliability of the CEMAUFL scale through a mixed-method research design, making it a valuable tool for researchers and educational institutions involved in mobile application development and language learning. Additionally, the article aims to contribute to the existing literature on MALL and promote a better understanding of the role of mobile applications in children's language education, especially in the context of Open and Distance Learning (ODL) during the COVID-19 pandemic.

2 A brief review of the CEMAUFL

The CEMAUFL was developed for use with 4th-grade students who are ten years old. The scale consists of 3 sections and 41 items. The first section includes the gender of the participants, whether they have a stable internet connection, whether they have mobile devices, the number of mobile devices, if any, the average daily use of mobile devices, whether they use a mobile application to learn English, and if the answer is yes, the names of the mobile applications.

In the second part, fourth-grade students can respond with Yes or No to statements about their use of mobile applications for educational purposes.

The third part contains expressions to assess the dimensions of mobile application use (such as the EBA mobile application, Dyned, Duolingo, and Memrise, which can be downloaded and used on smartphones and tablets) developed for foreign language teaching for 4th-grade students. The third part of the scale was designed to determine the use of mobile applications by 4th-grade students for learning English. It uses a 5-point Likert-type scale, with options such as "1 Strongly Disagree", "2 Disagree", "3 Undecided", "4 Agree", and "5 Strongly Agree". Additionally, a control item was included to identify if students answered the items randomly. The 5^{th} item of the scale represents the negative form of the 13^{th} item.

5^{th} item: I do not thinkmobile applications improve my success in English class.

13^{th} item: Mobile applications improve my success in English class.

Since it is a control item, the 5th item was not included in the factor analysis. The scale's highest score is 80, and the lowest is 16. Thus, the final version of the scale, consisting of 16 items, was used for validity and reliability studies. The first two sections of the scale serve as a descriptive framework for children's mobile application use for educational purposes; therefore, validity and reliability studies were not conducted for these sections. Since inferential statistics can be generated from the data obtained in the third section of the scale, validity and reliability studies were conducted only for this section.

2.1 The theoretical framework of the CEMAUFL

Scale development studies usually use experimental or theoretical procedures (Torgerson, 1958). The first step in the scale development process is establishing the conceptual framework

(Cohen & Swerdlik, 2012; Develis, 2012). Studies on children's foreign language learning, autonomy in foreign language learning, mobile learning, and MALL were reviewed in the literature. The theoretical framework of CEMAUFL was based on these studies.

In the Socio-Educational Model developed by Gardner (1985), one of the theories explaining foreign language acquisition, the importance of motivation and ability in language acquisition is emphasized. Gardner states that highly motivated individuals are primarily motivated to learn a foreign language and have positive attitudes toward learning (Masgoret & Gardner, 2003). Furthermore, the theory mentions that highly motivated individuals to learn a foreign language are more successful in learning the target language (Masgoret & Gardner, 2003). The factor willingness to use mobile applications and the related items in the third part of the scale was developed based on the Socio-Educational Model.

The Monitor Theory, which explains the process of foreign language acquisition, mentions that language is acquired. Children acquire a foreign language through meaningful input and meaning-oriented interaction (Karshen, 1978). Therefore, the emphasis in foreign language teaching is on skill acquisition. Javed et al. (2013) mention that language skills are divided into four categories: listening, reading, speaking, and writing. However, in addition to the four basic language skills (reading, writing, speaking, and listening), vocabulary, grammar, and pronunciation concepts are also used in foreign language teaching. These studies were the basis for developing the contribution to foreign language learning factor and its items.

Mobile applications are about learner autonomy. Learner autonomy is involved when learners download and start using a mobile application to learn English on their mobile devices. For this reason, studies on autonomy in foreign language learning have been focused on. The Mobile-Assisted Language Learning Through Learning Analytics for Self-Regulated Learning (MALLAS) model, developed by Viberg et al. (2020), explains MALL and provides a conceptual framework that captures the dimensions of self-regulated learning required for MALL. The factor of perceived control and its items in the third part of the scale were based on MALLAS.

After establishing the theoretical framework of the scale, an item pool was created. There are two ways to identify appropriate items: Deductive and Inductive methods (Boateng et al., 2018). The deductive method is conducted through a literature review and evaluation of scales in the field. On the other hand, the inductive method involves creating items from individuals' responses (Boateng et al., 2018; Hinkin, 1995). Direct observations and qualitative data, such as focus groups and individual interviews, can describe items inductively. The CEMAUFL scale item pool was created through deductive management: Reviewing the literature and evaluating scales in the field. A summary of the scales used in the development of the CEMAUFL and CEMAUFL's theoretical framework can be found in Table 1.

Table 1 Developing process of CEMAUFL items

Factor and items	The process of development of the item pool
Willingness to use mobile applications and their items in the third part of the scale	The Socio-Educational Model (Gardner, 1985) Attitude Scale Towards Mobile Learning (Demir & Akpınar, 2016) Mobile Learning Tools Acceptance Scale (Özer, 2017)
Contribution to foreign language learning and its items in the third part of the scale	Monitor Theory (Krashen, 1978) Mobile Learning Tools Acceptance Scale (Özer, 2017)
Perceived control and its items in the third part of the scale	MALLAS (Viberg, Watson & Kukulska-Hulme, 2020) Mobile Learning Attitude Scale (Önal & Tanık Önal, 2019)

The literature search yielded relevant studies, and from these, a pool of 43 items was generated. Initially, these items underwent language, expression, and wording corrections. Among them, 24 items were related to children's general use of mobile applications, while the remaining 19 were designed to assess children's usage of mobile applications for language learning. After careful review for semantic coherence and appropriateness, the number of items was reduced to 16. The factors for these 16 items were drawn from the existing literature, resulting in seven items falling under the "contribution to foreign language learning" factor, six items under the "willingness to use mobile applications" factor, and three items under the "perceived control" factor. A control (5^{th}) item was also incorporated to identify random or purposeful responses. As a result, the final version of the scale was developed, comprising a 3-factor structure with 17 items.

3 Methods

The CEMAUFL scale development process follows an exploratory sequential mixed-method research design. Qualitative methodology was used in the first stage of the scale development

process, while quantitative methodology was used in the second stage. The CEMAUFL's factors and items were derived from the literature and further refined through input from experts in the field. The scale development process can be listed as follows: the creation of an item pool, determination of appearance and content validity, pilot applications, validity, and reliability studies. The development process of the scale is shown in Table 2.

Table 2 The scale development process

Stage	Purpose	Data Collection Method	Sample	Analysis Method
Qualitative Stage	Theoretical Framework	Literature review		
	Item Pool	Literature review and opinions of experts		Content Analysis
	Appearance and Content Validity	Field Experts Assessment Scale	19 experts	Mean and Content Analysis
	Pilot Applications	CEMAUFL	37 students in the 4^{th} grade	Content Analysis
Quantitative Stage	Validity	CEMAUFL	309 students in the 4^{th} grade	Confirmatory Factor Analysis
	Reliability	CEMAUFL		Cronbach Alpha (α)

3.1 Study group

Experts who were pursuing their doctorates in ODL and working as English teachers carried out the appearance and content validity. Field experts were determined using the convenience sampling method. In this method, the researcher defines the study group, starting with the most accessible respondents until they reach a group of the size they need (Büyüköztürk et al., 2008). The determination of the study group was based on the volunteerism of the participants. The Field Experts Assessment Scale was sent to 21 experts selected to evaluate the scale, and 20 responded. The scale was excluded from the research sample since an expert evaluation form was not completed. Thus, feedback from 19 field experts was considered for appearance and content validity.

Pilot implementations of the CEMAUFL scale were carried out with 37 4th-grade students in a primary school. For validity and reliability studies, data were collected from 345 4th-grade students. However, since the scale form of 12 participants was empty, the data of these participants were excluded from the sample. Therefore, validity and reliability analysis was started with the data collected from 333 study groups.

3.2 Data collection tools

The Field Experts Assessment Scale was used to determine the appearance and content validity of the CEMAUFL. In the evaluation form, the participants were asked to examine the items belonging to the CEMAUFL and to evaluate the degree of suitability for the purpose to be measured from 1 to 10 (1 is the lowest - 10 is the highest). In the end, an explanation area was left for the participants, who were asked to indicate their opinions and suggestions about the suitability of the CEMAUFL items. CEMAUFL was used to determine the validity and reliability.

3.3 Data collection

The Field Experts Assessment Scale form was converted to Google Forms, and the generated link was electronically shared to assess the appearance and content validity. Two pilot studies were conducted using the CEMAUFL scale. The scale was adapted to Google Forms for the pilot studies, and the link was distributed to the students. With parental supervision, the students completed the scale, which included a section for participants to share their experiences, opinions, and suggestions.

To assess the validity and reliability of the CEMAUFL, the scale was administered to 4th-grade students. The scale was transferred to Google Forms for administration, and the link was electronically distributed. Under the supervision of their parents, the students completed the scale. The message sent along with the link is provided below:

"Dear Parents,

I hope this message finds you well. As part of my doctoral thesis, I am researching the psychometric properties of the Children's Educational Mobile Application Usage for Foreign Languages Scale. This study aims to understand better children's usage of mobile applications for language learning.

Please participate in this research by allowing your child to complete the scale through the provided link. Your child's answers will be used solely for scientific purposes and will keep all responses confidential.

Please contact me with any questions or concerns regarding the research or the scale.

Thank you for your time and participation."

3.4 Data analysis

The data collected by the field experts on the appearance and content validity of the scale items were analyzed and interpreted with the mean ($\overline{X}$). In addition, the opinions and suggestions about the scale in the explanatory section were subjected to content analysis.

During the pilot applications, participants' experiences with the scale were evaluated using content analysis, a qualitative analysis method, and direct quotes were shared.

CEMAUFL items and factors were identified from the literature and expert opinion. In other words, there is a predetermined structure. The predetermined construct was tested using Confirmatory Factor Analysis (CFA). While Exploratory Factor Analysis (EFA) is used for theory development, CFA aims to test an existing theoretical structure (Rennie, 1997). CFA is a highly advanced technique based on testing theories about hidden variables (Tabachnick & Fidell, 2013). DFA is an analysis in which a previously defined and restricted structure is tested to determine whether it is validated as a model. Within the scope of the reliability study of the CEMAUFL, the Cronbach Alpha (α) coefficient was examined to test the consistency of the scale items.

4 Findings

4.1 Appearance and content validity

The scale aims to determine the use of children's mobile applications to learn foreign languages. The 17 items in the third part of the scale, determined using the literature review, were categorized into three factors. The scale items were assigned to the factors of contribution to foreign language learning (7 items), willingness to use mobile applications (6 items), and perceived control (3 items). Participants rated the appropriateness of the scale items and factors for the scale from 1 to 10. They also wrote their opinions about the scale in the explanatory section.

Table 3 The averages of the expert assessment of the scale

	Willingness To Use Mobile Applications ($\overline{X}$)	Contribution To Foreign Language Learning ($\overline{X}$)	Perceived Control ($\overline{X}$)
	M1 = 9.1	Y1 = 8.10	A1 = 8.5
	M2 = 8.3	Y2 = 9.30	A2 = 7.6
	M3 = 7.7	Y3 = 9.10	A3 = 8.3
	M4 = 8.2	Y4 = 9.00	
	M5 = 8.7	Y5 = 9.00	
	M6 = 8.3	Y6 = 7.00	
		Y7 = 9.05	
Total Average	8.3	8.65	8.1

When Table 3 is viewed, the averages of the evaluations made by the field experts about the suitability of the scale items for the purpose can be seen. For example, when the average scores of the items are looked at, it is seen that the items with an average score of 9 and above are the first item for the willingness to use mobile applications factor and the second, third, fourth, fifth, and seventh items for the contributing to foreign language learning factor. Therefore, the items are considered appropriate for the purpose intended to be measured by the experts in the field.

Items with an average score of 8 and above are the 2^{nd}, 4^{th}, 5^{th}, and 6^{th} items for the willingness to use mobile applications factor; the first item for the contribution to foreign language learning factor; 1^{st} and 3^{rd} items for the perceived control factor. Therefore, these items are also suitable for being measured. However, one of the participants stated in the explanatory section that the third item under the perceived control factor should be simplified and made more understandable. Therefore, the item was rewritten, considering the recommendations of field experts.

Items with an average score of 7 and above are the 3^{rd} item for the willingness to use mobile applications factor, the 6th item for the contribution to foreign language learning factor, and the 2^{nd} item for the perceived control factor. One of the participants stated that the third item, under the willingness to use mobile applications factor, is more suitable for contributing to the foreign language learning factor. The necessary changes were made considering the average score and the recommendations of field experts. The participants also stated that the sentence should be revised about the second item in the perceived control factor. The item has been rewritten based on the average score and the opinions of field experts. Thus, the final version of the scale form was determined before the pilot applications.

4.2 Pilot applications

Two pilot implementations have been carried out for this scale. First, the scale was transferred to Google Forms for the pilot applications, and the link was sent to the students. The students completed the scale under the supervision of their parents. In addition, a description section has been added to the end of the scale so that participants can write their experiences, opinions, and suggestions about the scale.

The explanations from some students who made feedback about their experiences while answering the scale are as follows:

"I had no difficulty and did it in 5-6 minutes". ***Ece***

"It did not take even 5 minutes; I clicked on the link and marked it according to the questions". ***Ceren***

"I did not have a problem doing the survey; I was just confused; it took 5-10 minutes". ***Irmak***

"I did not have much difficulty, and I did it in 10 minutes". ***Berke***

The students completed the scale in an average of 5-6 minutes. As a result of the pilot application, all items are functional and understandable. Furthermore, it was observed that students were able to handle difficulties.

After the first pilot implementation, a second opinion was obtained from the field experts about the scale, and changes were made to the scale. The field experts' opinions led to amendments in the scale's instructions. To provide students with a clearer understanding of mobile applications, examples of mobile applications that support foreign language education were given in the instructions. Separate instructions were added for parents and students. Additionally, based on the opinion that the second and fifth items of the scale were similar, the second item was removed from the scale. As pronunciation is one of the essential skills in foreign language learning, an item about pronunciation was added. Furthermore, the seventeenth item was revised to use language and expressions that students can understand.

Due to the changes made, a second pilot application was carried out with 14 students studying in the fourth grade. The scale was transferred to Google Forms for the pilot implementation, and the link was shared with the students. The students completed the scale under the supervision of their parents. Additionally, a description section was added at the end of the scale for participants to share their experiences, opinions, and suggestions about the scale.

In the second pilot implementation, the participants' experiences related to the scale were analyzed using qualitative content analysis. The statements from the students who provided feedback about their experiences are as follows:

"I finished in 3 minutes; there was no place I did not understand in the questions," ***Yağmur***

"I completed it in 3-4 minutes, and there was no place I did not understand," ***Meryem***

"I completed it in 4 minutes; nothing happened that I did not understand," ***Bekir said.***

"I completed the survey in an average of 3-4 minutes. I had no difficulty understanding the questions. "I use these types of programs" ***Haktan.***

Telephone interviews were also conducted with 3 of the participants. Modifications were implemented on the scale, considering the feedback and suggestions from the participants.

4.3 Validity of the scale

After the pilot implementations, Confirmatory Factor Analysis (CFA) was performed to determine the validity of the updated scale. Before proceeding to CFA applications, it is necessary to question some basic concepts and possible situations to minimize the problems that may arise (Çokluk et al., 2018). This direction determined sample size, missing values, normality, straightness, multiple connection and singularity, and outliers. Calculations of the appropriateness of the data for factor analysis were made using SPSS.

4.3.1 Missing values

Data were collected from 345 students in the fourth grade for validity and reliability studies. However, since the scale form of the 12 participants was empty, the data belonging to these participants were included in the sample. Therefore, validity and reliability analyses were conducted using data collected from 333 participants.

4.3.2 Outliers

Before proceeding to statistical analyses, the outliers in the data set were determined. Factor analysis is sensitive to extreme values (Pallant, 2012). For this reason, whether there is an outlier in the data should be checked. Outliers are data that take values very far from the average of a particular distribution (Jarrell, 1994).

Univariate outliers can be determined with the help of some statistical methods, such as converting all scores in the distribution to standard scores. For this purpose, raw scores can be converted to standard z-scores (Çokluk et al., 2018, s. 13). In large samples ($n > 100$), the

z-score range is – 4 and +4 (Mertler & Vannatta, 2000). The sum of the scores given to the scale by the participants was between -4 and +4 when the z scores were examined.

Multiple outliers mean unusual combinations of scores for two or more variables. Each item acts as a variable, so that that factor analysis will be performed for 16 variables. Multiple outliers are more challenging to resolve and difficult to identify using the previously mentioned techniques. However, it can be determined by a statistical process known as the Mahalanobis distance (Mertler & Vannatta, 2005). The critical value for the Mahalanobis distance 16 variables is between 39.252 ($p < 0.001$) (Tabachnick & Fidell, 2013, p. 952). Outliers increase the error variance and reduce the power of statistical tests (Rasmussen, 1988). Pallant (2012, p. 58) states that deleting these values from the data set can solve outliers. After examining the data set of the current study in terms of outliers, it was decided to exclude the data of 24 participants from the analysis. In this case, the number of participants decreased from 333 to 309.

4.3.3 Sample size

An application was carried out with the fourth-grade students to perform factor analysis. Pallant (2012) states that at least 5 participants should be assigned to each item in the factor analysis. Another criterion for testing the suitability of the data structure for CFA in terms of sample size is the results of the Kaiser Mayer Olkin (KMO) test (Çokluk et al., 2018). Within the scope of this study, the data collected from 309 participants were used for CFA. Table 4 shows that the sample size of the data collected for CFA is excellent.

Table 4 Bartlett Sphericity test and KMO

KMO	χ^2	df	p
0.950	3418.797	120	0.000

4.3.4 Normality

In factor analysis, each item acts as a variable. Therefore, CFA is performed for 16 variables in the 3^{rd} part of the CEMAUFL scale. Multivariate normality is the normal distribution of all variables and all linear combinations of variables (Tabachnick & Fidell, 2013). The multivariate normality of the data is calculated with the Bartlett Sphericity Test (Cokluk et al., 2018). When Table 4 is examined, it is seen that the Bartlett test is significant ($p < 0.05$). This finding reveals that the criterion of multiple normality was met.

To examine the normality, one of the assumptions of the factor analysis, the distribution graphs of each item were examined. Histogram graphs gave images close to the normal distribution. In addition, the kurtosis and skewness values of each item were examined. Fabrigar et al. (1999, p. 283) state that the skewness value should not be greater than two and the kurtosis value should not be greater than 7 when the maximum similarity technique is used in factor analysis. Thus, upon examining the skewness and kurtosis values for each item, it can be concluded that the data set meets the univariate normality assumption.

4.3.5 Linearity

Linearity can be defined as having a correct relationship between two variables. The linear correlation coefficient (r) finds a linear relationship between two variables. The decision point for linearity varies according to the number of people in the sample. For n = 100, the decision is 0.196 (-1 to -0.196 and +0.196 to +1) (Cokluk et al., 2018). When the correlation coefficients of the data set were examined, it was concluded that there was a linear relationship between the variables.

4.3.6 Multicollinearity and singularity

For us to perform CFA, there should not be a strong relationship between the variables. Multicollinearity is when test items are highly correlated in pairs. Therefore, tolerance and VIF values should be examined to determine multicollinearity. The tolerance values were between 0.224 and 0.712, and the VIF values were between 1.405 and 4.62. The fact that the tolerance values are more significant than 0.10, and the VIF values are less than 10 (Akbulut, 2010) shows no multicollinearity and singularity problem in the data set. The data are suitable for factor analysis. As a result, it was determined that the data set for factor analysis was appropriate.

4.3.7 Confirmatory factor analysis

The items and variables of the scale were determined using the related literature and with expert opinions. Since it is a predetermined structure, CFA has been performed. The factors of the scale and the items of the factors are given in Table 5.

Table 5 Factors of the scale and the items of the factors

Factors	Number of items	Items
Willingness to use mobile applications	4	1, 2, 3, 4
Contribution to foreign language learning	10	6, 7, 8, 9, 10, 11, 12, 13, 14
Perceived control	3	15, 16, 17

CFA of the scale was carried out with the Jamovi program, an open-source application. The data were transferred to the SPSS program and then to the Jamovi. One of the most frequently used techniques as an estimation method, the maximum likelihood method, was used.

Factor loads indicate the weight or burden of representing the latent variable in the observed variable (Cokluk et al., 2018). In Table 6, factor loadings of the model are given.

Table 6 Factor loads of the scale model

Items		Factor	Estimate	S.E.	Z	p
Item 1	<—	Willingness to use mobile apps	0.790	0.0518	15.25	< 0.001
Item 2	<—	Willingness to use mobile apps	0.669	0.0584	11.45	< 0.001
Item 3	<—	Willingness to use mobile apps	0.567	0.0561	10.12	< 0.001
Item 4	<—	Willingness to use mobile apps	0.723	0.0541	13.38	< 0.001
Item 6	<—	Contribution to foreign language learning	0.799	0.0555	14.40	< 0.001
Item 7	<—	Contribution to foreign language learning	0.825	0.0415	19.88	< 0.001
Item 8	<—	Contribution to foreign language learning	0.827	0.0445	18.56	< 0.001
Item 9	<—	Contribution to foreign language learning	0.829	0.0455	18.24	< 0.001
Item 10	<—	Contribution to foreign language learning	0.845	0.0479	17.64	< 0.001
Item 11	<—	Contribution to foreign language learning	0.789	0.0479	16.47	< 0.001
Item 12	<—	Contribution to foreign language learning	0.797	0.0464	17.15	< 0.001
Item 13	<—	Contribution to foreign language learning	0.841	0.0459	18.32	< 0.001
Item 14	<—	Contribution to foreign language learning	0.761	0.0491	15.51	< 0.001
Item 15	<—	Perceived control	0.718	0.0531	13.53	< 0.001
Item 16	<—	Perceived control	0.803	0.0477	16.83	< 0.001
Item 17	<—	Perceived control	0.573	0.0624	9.18	< 0.001

As seen in Table 6 above, it is seen that all paths in the model of the CEMAUFL are statistically significant ($p < 0.005$). Estimates (estimates) values indicate a strong relationship between the latent and observed variables.

The model of CFA carried out to examine the theoretical model of the scale is given in Figure 1.

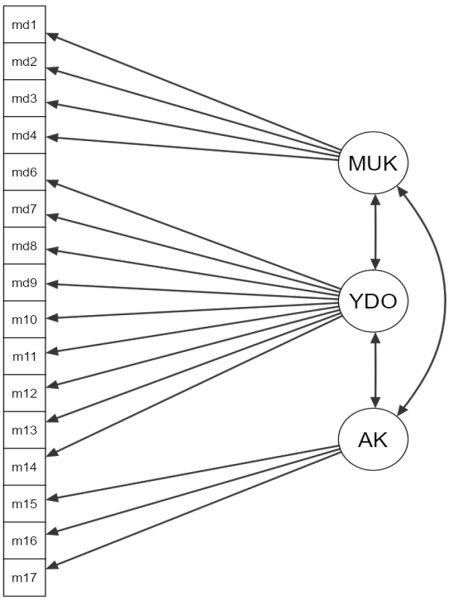

Figure 1 Confirmatory factor analysis model (MUK: Willingness to Use Mobile Applications, YDO: Contribution to Foreign Language Learning, AK: Perceived Control)

Chi-Square (χ^2), Comparative Fit Index (CFI), and Root Mean of Square Error of Approximation (RMSEA) statistics were calculated to determine whether the data and the original structure fit. Table 7 shows χ^2, CFI, and RMSEA values.

Table 7 RMSEA, CFI, and χ^2 statistics of the scale

n	χ^2	df	χ^2/df	CFI	RMSEA
309	257	101	2.54	0.954	0.07

Traditionally, the most frequently cited fit index in research reports is the Chi-square statistic (χ^2) (Munro, 2005). The p-value of the χ^2 statistic is affected by the sample size and may result in the model being rejected unless there are huge samples (Waltz et al., 2010). In this case, the value obtained by dividing the χ^2 value by the degrees of freedom is considered (Şimşek, 2007). χ^2 and degrees of freedom (df) were obtained with the Jamovi program's help. The χ^2/df ratio is 2.54. A value of 5 or less indicates that the model and data fit well (Gillaspy, 1996). This finding shows that the proposed model for the scale fits with the collected data.

CFI gives the fit of the model by comparing it with a model that assumes no relationship between the variables (Cokluk et al., 2018). The CFI value ranges from 0 to 1. CFI exceeding 0.90 is taken as a good model indicator. When Table 7 is examined, it is seen that the CFI value is 0.954. Since the value is close to 1, it can be stated that the model is perfect.

RMSEA is an index used to estimate population variance. This index takes a value between 0 and 1. Therefore, RMSEA being close to 0 indicates no difference between the population and sample (Brown, 2016). Therefore, in Table 7, the sample of the data collected for the scale represents the universe.

4.4 Reliability of the scale

Reliability is measured to determine whether all items in a measurement tool serve the same purpose and whether the items form a whole (Akbulut, 2010). Internal consistency was calculated in the Jamovi using the data obtained from the application to determine the scale's factor analysis. The data regarding the mean, standard deviation, and Cronbach Alpha value of the scale are given in Table 8.

Table 8 Mean, standard deviation, and alpha

Number of items	n	$\overline{X}$	σ	α
16	309	3.68	0.750	0.945

Table 8 shows the number of participants (n), mean ($\overline{X}$), standard deviation (σ), and Cronbach Alpha value (α) related to 16 items of the scale. When the data collected from 309 participants are examined, it is seen that the mean of the scale is 3.68 out of 5 points. The standard deviation is 0.750. The standard deviation is the average of the distances of all individuals in a sequence from the mean (Akbulut, 2010). Individuals get closer to the mean as this value decreases, and a homogeneous distribution is observed (Akbulut, 2010). When the data set's standard deviation is examined, it can be stated that the distribution is normal. It is seen that the Cronbach Alpha value is 0.945. Since its reliability value is over 0.90, it can be stated that the

CEMAUFL is a very reliable measurement tool (Pallant, 2012). In addition, as seen in Table 9, when an item is deleted, what kind of change might occur in the scale's Cronbach Alpha (α) value was also examined.

When Table 9 is examined, the reliability coefficient of the scale is seen when each item is removed. For example, when one of the scale items is removed, it is seen that the Cronbach alpha value of the scale is close to 0.945. Therefore, all items were kept in the scale.

5 Conclusion and recommendations

In conclusion, this study successfully developed the CEMAUFL scale, a valuable tool for assessing children's usage of mobile applications to support foreign language learning. The scale demonstrated validity and reliability through rigorous psychometric testing, including Confirmatory Factor Analysis (CFA) and reliability studies, establishing it as a significant contribution to the literature. Being one of the first psychometric scales in this area, CEMAUFL fills a crucial gap and offers researchers and educators a reliable and valid instrument to explore the impact of mobile applications on language learning in young students.

Given the increasing significance of ODL during the Covid-19 pandemic, CEMAUFL assumes even more importance as a means of understanding and maximizing the potential benefits of mobile applications in education. It presents opportunities for further research in the field, particularly when investigating the relationship between children's educational mobile application use and other variables. Moreover, the scale's applicability for institutions developing or updating mobile applications makes it a valuable resource for educational practitioners and developers.

Table 9 Items total statistics

Items	The original version of items	Cronbach's α
Mobile applications make learning English enjoyable.	Mobil uygulamalar İngilizce öğrenmemi zevkli hale getirir.	0.942
When learning English, I prefer to learn using mobile applications.	İngilizce öğrenirken mobil uygulamalar kullanarak öğrenmeyi tercih ederim	0.945
If my teacher allows, I would like to use mobile applications in English class.	Öğretmenimin izin vermesi halinde, İngilizce dersinde mobil uygulama kullanmak isterim.	0.946
I look forward to the times when I use mobile applications while learning English.	İngilizce öğrenirken mobil uygulama kullandığım zamanları sabırsızlıkla beklerim	0.943
Mobile applications increase my writing skills in English	Mobil uygulamalar İngilizce yazma becerimi arttırır.	0.939
Mobile applications increase my English vocabulary.	Mobil uygulamalar İngilizce kelime bilgimi arttırır.	0.942
Mobile applications improve my English grammar.	Mobil uygulamalar İngilizce dil bilgimi arttırır.	0.939
Mobile applications enhance my reading skills in English.	Mobil uygulamalar İngilizce okuma becerimi arttırır.	0.939
Mobile applications enhance my speaking skills in English.	Mobil uygulamalar İngilizce konuşma becerimi arttırır.	0.940
Mobile applications improve my listening skills in English.	Mobil uygulamalar İngilizce dinleme becerimi arttırır.	0.940
Mobile applications help me learn how to pronounce English words.	Mobil uygulamalar İngilizce kelimelerin nasıl söyleneceği konusunda bana fayda sağlar.	0.940
Mobile applications improve my success in English class.	Mobil uygulamalar İngilizce dersindeki başarımı arttırır.	0.939
I easily remember the English language I learned from mobile applications.	Mobil uygulamalardan öğrendiğim İngilizceyi kolay hatırlarım.	0.941
I easily learn how to use new mobile applications that support learning English.	İngilizce ile ilgili yeni mobil uygulamaları nasıl kullanacağımı kolayca öğrenirim.	0.943
I am confident in learning English with mobile applications.	Mobil uygulamalarla İngilizce öğrenme konusunda kendime güvenirim.	0.941
I can learn English by myself using mobile applications.	Mobil uygulamalarla kendi kendime İngilizce öğrenebilirim.	0.947

The scale has been designed to assess children's educational mobile application usage for foreign language learning in a comprehensive and language-neutral manner; it can be adaptable to various languages. Moreover, the scale can evaluate mobile applications that support foreign language learning for children across different language contexts. It can be applied to various mobile applications targeting language learning in various languages. However, it may require further validation and testing with diverse language learners and app users to ensure the scale's validity and reliability when applied to different languages and mobile applications.

It is essential to acknowledge the limitations of this study and the CEMAUFL scale. The research sample was limited to 365 participants, and data on children's mobile application use were based on personal statements. Additionally, the assumption that 4th-grade students completed the scale with their parents may be a potential limitation.

In conclusion, the CEMAUFL scale opens new avenues for research, providing valuable insights into the effectiveness of mobile applications in supporting foreign language learning for young learners. This scale will encourage further investigation in the field and enhance language education through innovative technology-driven approaches.

Acknowledgements

This article was produced from the first writer's doctoral thesis titled "The Effects of Current Mobile Applications Used in Foreign Language Teaching on 4th Grade Students' Success in English Class," conducted under the supervision of a second writer.

Conflicts of interest

The authors declare that they have no conflict of interest.

References

Akbulut, Y. (2010). Sosyal bilimlerde SPSS uygulamaları. İstanbul: İdeal Kültür Yayıncılık.

Albright, J. J., & Park, H. M. (2009). Confirmatory factor analysis using Amos, Lisrel, Mplus, and SAS/STAT. Bloomington: University Information Technology Services Center for Statistical and Mathematical Computing Indiana University Press.

Alkan, C. (2011). Eğitim Teknolojisi (8. Baskı). Ankara: Anı Yayıncılık.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioural research: A primer. Frontiers in Public Health, 6, 149. https://doi.org/10.3389/fpubh.2018.00149

Booton, S. A., Hodgkiss, A., & Murphy, V. A. (2023). The impact of mobile application features on children's language and literacy learning: a systematic review. Computer Assisted Language Learning, 36(3), 400-429. https://doi.org/10.1080/09588221.2021.1930057

Brown, T. H., & Mbati, L. S. (2015). Mobile learning: Moving past the myths and embracing the opportunities. International Review of Research in Open and Distributed Learning, 16(2), 115-135. https://doi.org/10.19173/irrodl.v16i2.2071

Büyüköztürk, Ş., Çakmak, E. K., Akgün, Ö. E., Karadeniz, Ş., & Demirel, F. (2008). Bilimsel araştırma yöntemleri. Ankara: Pegem Akademi.

Cohen, R. J., & Swerdlik, M. (2002). Psychological testing and assessment: An introduction to test and measurement. Boston: McGraw-Hill.

Çam, E., & Uysal, M. (2017). Mobil uygulamaların eğitsel amaçlı kullanımı: Bir ölçek geliştirme çalışması. Uluslararası Türk Eğitim Bilimleri Dergisi, 9, 559-567.

Can, E. (2020). Coronavirüs (Covid-19) pandemisi ve pedagojik yansımaları: Türkiye'de açık ve uzaktan eğitim uygulamaları. AUAd, 6(2), 11-53.

Çokluk, Ö., Şekercioğlu, G., & Büyüköztürk, Ş. (2012). Sosyal bilimler için çok değişkenli istatistik: SPSS ve LISREL uygulamaları (Vol. 2). Ankara: Pegem Akademi.

Demir, K., & Akpınar, E. (2016). Mobil öğrenmeye yönelik tutum ölçeği geliştirme çalışması. Eğitim Teknolojisi Kuram ve Uygulama, 6, 59-79. https://doi.org/10.17943/etku.83341

Dockrell, J. E., & Marshall, C. R. (2015). Measurement issues: Assessing language skills in young children. Child and Adolescent Mental Health, 20(2), 116-125. https://doi.org/10.1111/camh.12072

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. Psychological Methods, 4(3), 272-279. https://doi.org/10.1037/1082-989X.4.3.272

Fernandes, T., & Pedroso, R. (2017). The effect of self-checkout quality on customer satisfaction and patronage in a retail context. Service Business, 11, 69-92. https://doi.org/10.1007/s11628-016-0302-9

Fırat, M. (2019). Uygulamadan kurama açık ve uzaktan öğrenme. Ankara: Nobel Yayınları.

Gardner, R. C. (1985). Social psychology and second language learning: The role of attitudes and motivation. London: Edward Arnold.

Gillaspy, J. A. (1996). A primer on confirmatory factor analysis. Paper presented at Annual Meeting of the Southwestern Educational Research Association. Austin, Texas. New Orleans, LA.

Hinkin, T. R. (1995). A review of scale development practices in the study of organizations. Journal of Management, 21, 967-988. https://doi.org/10.1177/014920639502100509

Jarrell, M. G. (1994). A comparison of two procedures, the Mahalanobis Distance and the Andrews-Pregibon Statistic, for identifying multivariate outliers. Research in the Schools, 1, 49-58.

Javed, M., Juan, W. X., & Nazli, S. (2013). A study of students' assessment in writing skills of the English language. International Journal of Instruction, 6(2), 129-144.

Johnson, J. S., & Newport, E. L. (1991). Critical period effects on universal properties of language: The status of subjacency in the acquisition of a second language. Cognition, 39, 215-258. https://doi.org/10.1016/0010-0277(91)90054-8

Kalaycı, Ş. (2006). Faktör analizi. SPSS uygulamalı çok değişkenli istatistik teknikleri. Asil Yayın Dağıtım.

Kline, P. (1993). A Handbook of Psychological Testing (2nd Ed.). London: Routledge; Taylor & Francis Group.

Knowles, M. S. (1975). Self-directed learning: A guide for learners and teachers. New York: Association Press.

Krashen, S. (1978). Individual variation in the use of the Monitor. In W. Ritchie (Ed.), Principles of Second Language Learning (pp. 175-183). New York: Academic Press.

Masgoret, A. M., & Gardner, R. C. (2003). Attitudes, motivation, and second language learning: A meta-analysis of studies conducted by Gardner and Associates. Language Learning, 53(1), 123-163. https://doi.org/10.1111/1467-9922.00212

Mertler, C. A., & Vannatta, R. A. (2005). Advanced and multivariate statistical methods. California: Pyrczak.

Munajah, R., Sumantri, M. S., & Yufiarti, Y. (2022). The use of digital storytelling to improve students' writing skills. Advances in Mobile Learning Educational Research, 3(1), 579-585. https://doi.org/10.25082/AMLER.2023.01.006

Nunan, D. (2003). The impact of English as a global language on educational policies and practices in the Asia-Pasific region. TESOL Quarterly, 37(4), 589-613. https://doi.org/10.2307/3588214

Önal, N., & Önal Tanık, N. (2019). İngilizce mobil öğrenme tutum ölçeğinin Türkçe'ye uyarlanması: Geçerlik ve güvenirlik çalışması. Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi, 7(1), 283-289.

Önal, N., Çevik, K.K.& Şenol, V. (2022). The effect of SOS Table learning environment on mobile learning tools acceptance, motivation and mobile learning attitude in English language learning, Interactive Learning Environments, 30(5), 834-847. https://doi.org/10.1080/10494820.2019.1690529

Özer, Ö. (2017). Mobil destekli öğrenme çevresinin yabancı dil öğrencilerinin akademik başarılarına, mobil öğrenme araçlarını kabul düzeylerine ve bilişsel yüke etkisi. Yayınlanmamış Doktora Tezi. Mersin: Mersin Üniversitesi.

Pallant, J. (2012). SPSS survival manual: A step-by-step guide to data analysis using SPSS (Version 12). Australia: National Library of Australia Cataloguing-in-Publication.

Papadakis, S., & Kalogiannakis, M. (Eds.). (2022). STEM, robotics, mobile apps in early childhood and primary education: Technology to promote teaching and learning. Springer Nature. https://doi.org/10.1007/978-981-19-0568-1

Papadakis, A., Barianos, A., Kalogiannakis, M., Papadakis, S., & Vidakis, N. (2022). ARION: A Digital eLearning Educational Tool Library for Synchronization Composition & Orchestration of Learning Session Data. Applied Sciences, 12(17), 8722. https://doi.org/10.3390/app12178722

Papadakis, S., & Kalogiannakis, M. (Eds.). (2020). Handbook of research on using educational robotics to facilitate student learning. IGI Global. https://doi.org/10.4018/978-1-7998-6717-3

Papadakis, S., Alexandraki, F., & Zaranis, N. (2021). Greek parents' app choices and young children's smart mobile usage at home. In Interactive Mobile Communication, Technologies and Learning (pp. 39-50). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-96296-8_4

Papadakis, S., Kalogiannakis, M., & Gözüm, A. I. C. (2022). STEM, STEAM, computational thinking, and coding: Evidence-based research and practice in children's development. Frontiers in Psychology, 13, 1110476. https://doi.org/10.3389/fpsyg.2022.1110476

Putra, F. K. K., & Law, R. (2023). Critical success factors for virtual hotel operator partnership with small-and medium-sized hotels: perspectives of owners and operators. Journal of Hospitality and Tourism Insights, 1-21. https://doi.org/10.1108/JHTI-06-2022-0244

Queen, B., & Lewis, L. (2011). Distance education courses for public elementary and secondary school students: 2009-10 (NCES 2012-009). U.S. Department of Education, National Center for Education Statistics.

Rennie, K. M. (1997). Exploratory and confirmatory rotation strategies in exploratory factor analysis. Paper presented at the Southwest Educational Research Association, TX Annual Meeting: Austin.

Sato, M. Skill learning theories and language teaching: Different strokes for different folks. In Practice and Automatization in Second Language Research (pp. 63-86). Routledge. https://doi.org/10.4324/9781003414643-4

Schumacker, R. E., & Beyerlein, S. T. (2000). Confirmatory factor analysis with different correlation types and estimation methods. Structural Equation Modeling, 7(4), 629-636. https://doi.org/10.1207/S15328007SEM0704_6

Şencan, H. (2005). Sosyal ve davranışsal ölçümlerde güvenirlik ve geçerlik. Ankara: Seçkin Yayıncılık.

Şimşek, Ö. F. (2007). Yapısal eşitlik modellemesine giriş, temel ilkeler ve Lisrel uygulamaları. Ankara: Ekinoks.

Skaraki, E. (2021). Reinforcing preschoolers' phonemic awareness through the use of tablets. Advances in Mobile Learning Educational Research, 1(1), 28-36. https://doi.org/10.25082/AMLER.2021.01.004

Slavoff, G. R., & Johnson, J. S. (1995). The effects of age on the rate of learning a second language. Studies in Second Language Acquisition, 17, 1-16. https://doi.org/10.1017/S0272263100013723

Sözeri, V., & Harmanşah, C. (2018). İngilizce öğretimi için mobil uygulama örneği. MCBÜ Soma Meslek Yüksekokulu Teknik Bilgiler Dergisi, 27, 62-70.

Strataki, A. (2022). An evaluation of educational apps for preschool-age children in Android and iOS. Advances in Mobile Learning Educational Research, 2(1), 278-288. https://doi.org/10.25082/AMLER.2022.01.012

Sunar, P. K., Dahal, N., & Pant, B. (2022). Using digital stories during COVID-19 to enhance early-grade learners' language skills. Advances in Mobile Learning Educational Research, 3(1), 548-561. https://doi.org/10.25082/AMLER.2023.01.003

Sutrisni, D. M., Utaminingsih, S., Murtono, M., Mariam, I. O., & Pratama, H. (2022). The effectiveness of the android-based budiran game assisted by smart apps creator 3 to improve science learning outcomes of fourth graders in theme 1. Advances in Mobile Learning Educational Research, 2(2), 483-492. https://doi.org/10.25082/AMLER.2022.02.017

Tabachnick, B. G., & Fidell, L. S. (2013). Using Multivariate Statistics (3rd ed.). New York Harper Collins.

Torgerson, W. S. (1958). Theory and methods of scaling. New York: Wiley.

Torres, E. N. (2023). Why great service is difficult to achieve: Insights from theory and practice. International Journal of Contemporary Hospitality Management. https://doi.org/10.1108/IJCHM-05-2022-0580

Tsoukala, C. (2021). STEM integrated education and multimodal educational material. Advances in Mobile Learning Educational Research, 1(2), 96-113. https://doi.org/10.25082/AMLER.2021.02.005

Viberg, O., Wasson, B., & Kukulska-Hulme, A. (2020). Mobile-assisted language learning through learning analytics for self-regulated learning (MALLAS): A conceptual framework. Australasian Journal of Educational Technology, 36(6), 34-52. https://doi.org/10.14742/ajet.6494

Wang, X., Markert, C., & Sasangohar, F. (2023). Investigating popular mental health mobile application downloads and activity during the COVID-19 pandemic. Human Factors, 65(1), 50-61. https://doi.org/10.1177/0018720821998110

Wu, Q. (2015). Designing a smartphone app to teach English (L2) vocabulary. Computers & Education, 85, 170-179. https://doi.org/10.1016/j.compedu.2015.02.013

Yang, J. (2013). Mobile assisted language learning: A review of the recent applications of emerging mobile technologies. English Language Teaching, 6(7), 19-25. https://doi.org/10.5539/elt.v6n7p19

Yaşlıoğlu, M. M. (2017). Sosyal bilimlerde faktör analizi ve geçerlik: keşfedici ve doğrulayıcı faktör analizlerinin kullanılması. Istanbul University Journal of the School of Business, 46, 74-85.