The Khan Academy Platform in the Mathematics Learning Process in Military Training

Magaly Margarita Narváez Ríos
Social Sciences, Education and Humanities, Private Technical University of Loja
San Cayetano Alto, C. París, 110107, Loja, Ecuador
&
Faculty of Human Sciences and Education, Technical University of Ambato
Av. de Los Chasquis y Río Payamíno, 180329, Ambato, Ecuador
E-mail address: mm.narvaez@uta.edu.ec
ORCID: https://orcid.org/0000-0001-9431-4972

José Ramón Delgado Fernández
Social Sciences, Education and Humanities, Private Technical University of Loja
San Cayetano Alto, C. París, 110107, Loja, Ecuador
E-mail address: jrdelgado66@utpl.edu.ec
ORCID: https://orcid.org/0000-0002-9176-7666

Derling José Mendoza Velazco
Initial Education Career, National University of Education
Panamericana 1, Chuquipata, 030154, Cañar, Ecuador
E-mail address: derling.mendoza@unae.edu.ec
ORCID: https://orcid.org/0000-0001-8275-3687

Diego Santiago Andrade Naranjo
Faculty of Human Sciences and Education, Technical University of Ambato
Av. de Los Chasquis y Río Payamíno, 180329, Ambato, Ecuador
E-mail address: ds.andrade@uta.edu.ec
ORCID: https://orcid.org/0000-0002-5536-6267

ABSTRACT

Aim. To analyse the effectiveness of Khan Academy in strengthening mathematical learning among first-year cadets in an Ecuadorian military training context.

Methods. A quantitative quasi-experimental one-group pre-test/post-test design was conducted with 65 cadets from the Escuela de Formacion de Soldados "Vencedores del Cenepa". Instrument reliability was examined through Cronbach's alpha (0.834-

0.869). The main analysis relied on descriptive statistics and a paired-samples t-test, complemented by confidence intervals and effect size estimation.

Results. Mathematics performance increased significantly from the pre-test (M = 5.85, SD = 3.03) to the post-test (M = 16.30, SD = 1.57), t(64) = 27.39, $p < .001$, with a very large effect size (dz = 3.40). The mean gain was 10.45 points. Satisfaction with Khan Academy was high when treated as a continuous 10-item scale (M = 47.69 out of 50, SD = 3.00), although satisfaction was not significantly associated with score gain.

Conclusions. Khan Academy was associated with substantial improvement in mathematics performance in this military higher education setting. Nevertheless, the findings should be interpreted cautiously because the design lacked a control group and the pre-test and post-test were not strictly parallel forms.

Practical application. The study offers empirical support for the structured integration of digital platforms into mathematics teaching in specialised higher education and provides an institutional model that may inform curricular innovation in comparable contexts.

Keywords: mathematics education, Khan Academy, digital platforms, military training, Ecuador

Introduction

In recent decades, mathematics education has been recognised as a strategic pillar for the development of technological and digitised societies. Online learning systems such as Khan Academy, Coursera and EdX have enabled students from different continents to access structured content, interactive resources and adaptive assessments. In countries across Asia and Europe, multiple studies have documented that these digital platforms have strengthened autonomous learning, increased motivation, and offered alternative ways to overcome deficiencies in mathematical skills (Wang & Sun, 2025). In addition, international organisations such as UNESCO have pointed out that the integration of digital environments in science and mathematics education contributes to democratising access to knowledge and preparing students to face complex problems in globalised contexts (Wang & Sun, 2025). As a result, the adoption of digital platforms is no longer considered an isolated innovation, but an essential component of modern education systems.

In Latin America, significant progress has been made, although significant challenges remain. Several countries have promoted educational digitisation and teacher training programmes to incorporate technological tools into the classroom. In Mexico, Brazil, Chile, and Argentina, for example, national technological integration projects have been developed that included open platforms for teaching mathematics (Saracostti et al., 2025). Recent research has shown that these experiences improved conceptual understanding in algebra, geometry, and calculus, provided that the programmes were accompanied by teacher training and adequate technological infrastructure (Borchers & Cunha, 2025). However, the region still faces digital divides, inequalities in internet access, and shortcomings in the adaptation of digital resources to specific sociocultural contexts. These contrasts position Latin America as a living laboratory for the study of technological incorporation strategies, where successful experiences and structural limitations coexist (Zaidi et al., 2025).

In the case of Ecuador, national mathematics learning indicators have shown gaps compared to international averages. Reports from the National Institute for Educational Evaluation have highlighted recurring difficulties in logical reasoning, problem solving, and algebraic comprehension among secondary and higher education students (Tosi & Theobald, 2026). Although the Ministry of Education has promoted the use of digital platforms, implementation has been uneven and poorly documented in military training institutions. The Army Soldier Training School (ESFORSE) is a unique setting where mathematics education is essential for developing the analytical and strategic skills necessary for military life. However, until now, there have been no systematic studies exploring the impact of Khan Academy as a teaching resource in this environment. The absence of academic evidence on the relevance, effectiveness, and limitations of this tool in military training created a critical knowledge gap. If this problem were not addressed, the gap between the cognitive demands of the military profession and the limitations detected in mathematics learning could widen, compromising the quality of training for future soldiers.

Given this situation, the following central question emerged:

- How does the use of the Khan Academy platform affect the mathematics learning process of military training students at the Army Soldier Training School (ESFORSE)?

In line with the question posed, the study set the following objective:

- To analyse the use of the Khan Academy platform as a mathematics learning strategy for military training students at the Army Soldier Training School (ESFORSE).

Based on the contextual diagnosis, the following hypotheses were proposed:

- Null hypothesis (H_0): The use of the Khan Academy platform does not have a significant effect on mathematics learning among ESFORSE students.
- Alternative hypothesis (H_1): The use of the Khan Academy platform significantly improves mathematics learning among ESFORSE students.
- Theoretical Framework

Khan Academy as a Pedagogical Resource

Khan Academy was created in 2007 with the intention of offering clear and accessible explanations of complex mathematical topics. Since then, it has expanded to more than 190 countries, with millions of active users each month (Rueda-Gómez et al., 2023; Yu & Yang, 2024). Its features have made it a unique educational platform: short explanatory videos, interactive exercises, immediate feedback, gamification through badges, and personalised tracking of student progress.

Empirical studies reported that more than 80% of students who used Khan Academy in algebra or geometry courses perceived a positive effect on their learning (Egodage et al., 2025). The platform proved particularly useful in identifying knowledge gaps and allowing students to progress from simple to complex concepts at their own pace. In university contexts, it was also observed to promote self-confidence and a willingness to tackle more advanced mathematical content (Anggraini et al., 2026; Huang et al., 2025; Lemos et al., 2024).

Another key feature of Khan Academy is that it is free of charge. In education systems with budgetary constraints, this represents added value, as it facilitates universal access and removes economic barriers. The literature agrees that democratising access to mathematical knowledge through open platforms is an important step towards educational equity (Amalric et al., 2023; Fruett et al., 2024).

Khan Academy Instructional Components

The platform has four main components: interface, interactivity, navigation, and gamification. The interface is simple and user-friendly, allowing users to select content and adapt it to different levels of difficulty. Interactivity occurs through teacher mediation and peer interaction, which promotes the social construction of knowledge. Navigation offers flexible learning paths that each student can adjust according to their level of proficiency and interests. Finally, gamification introduces a system of badges and rewards that stimulates motivation and creates a more dynamic learning environment (Amalric et al., 2023; Traverso-Condori et al., 2024).

These components are linked to constructivist theories that understand learning as an active process. Students do not simply receive information, but participate in its construction, reflect on their mistakes, and adjust their strategies based on the feedback they receive (Saracostti et al., 2025). The literature also relates the design of Khan Academy to the principles of self-directed learning theory, where individuals take a leading role in their own learning process.

Didactics of Mathematics and Associated Theories

Teaching mathematics cannot be reduced to the transmission of formulas. It is a complex process in which logical reasoning, abstraction, and practical application converge. Theories such as Didactic Situations propose that learning occurs when students face a problem and seek solutions through interaction with their environment (Gleichmann et al., 2025). Teacher mediation is key to guiding reflection and consolidating knowledge. Another relevant perspective is the theory of social constructivism, which emphasises the importance of collaborative work and interaction with peers. Under this approach, digital platforms do not replace teachers but rather expand the possibilities for interaction and provide access to resources that enrich the learning process (Karimov et al., 2025).

Recent research suggests that the immediate feedback offered by Khan Academy reduces maths anxiety and allows for timely correction of errors (Li & Yin, 2025). Likewise, the integration of gamification elements promotes intrinsic motivation and persistence in the face of complex tasks. These aspects are fundamental in a field where students often experience high levels of frustration and dropout rates (Ngu et al., 2025).

Empirical Evidence About Khan Academy

Several studies confirmed the positive impact of Khan Academy in various educational contexts. In Argentina, its implementation in engineering courses allowed students to improve their problem-solving skills and validate their knowledge more reliably (Ochogboju & Díez-Palomar, 2025). In Mexico and Colombia, the systematic use of the platform increased grades in basic mathematics courses and reduced failure rates (Saracostti et al., 2025).

Benefits were also reported in the flipped classroom model. By assigning Khan Academy videos and exercises as homework, class time could be devoted to discussion and collaborative problem solving. This dynamic promoted deeper learning and greater interaction between students and teachers (Abreh et al., 2025).

In Peru, studies on student perseverance showed that the platform not only contributed to content learning but also to the development of self-regulation and autonomy skills (Tosi & Theobald, 2026). However, the literature warns that the effectiveness of the tool depends on pedagogical guidance and the institutional context.

National Context and Research Gap

In Ecuador, national education assessment reports pointed to persistent deficiencies in logical reasoning, algebra, and problem solving. Official programmes introduced digital resources, but empirical evidence on their impact in military institutions is virtually non-existent (Poesia & Goodman, 2023). The Army Soldier Training School (ESFORSE) is a unique space where mathematics plays a central role in academic and strategic training. However, to date, no research has been conducted to analyse the effect of Khan Academy in this context. The absence of previous studies creates a knowledge gap that this work seeks to fill. Exploring the impact of the platform on military training will provide a better understanding of its scope, limitations and potential to strengthen the mathematical learning of future soldiers (González-Rogado et al., 2025; Rueda-Gómez et al., 2023).

Method

Research Design and Approach

The study adopted a quantitative approach and a quasi-experimental one-group pre-test/post-test design. The purpose was to examine whether the use of Khan Academy as a supplementary instructional platform was associated with changes in cadets' mathematics performance in the topic of systems of linear equations. The design relied on repeated measurement of the same participants before and after the intervention, thereby allowing within-subject comparison of performance over time.

Participants and Context

The study was conducted at the Escuela de Formacion de Soldados "Vencedores del Cenepa" (ESFORSE), Ecuador. The population comprised of 65 first-year applicants enrolled in the CBOS (+) LAMINA CHIGUANO cohort. A non-probabilistic intentional sampling strategy was used because all students from the two available class groups ('F' and 'U') who completed the instructional process and both achievement measurements were included in the analysis.

Intervention

The instructional intervention focused on systems of linear equations and incorporated Khan Academy as a structured support resource. According to the original instructional plan, three videos were selected from 27 available videos and 15 short activities or workshops were selected from 37 available platform activities. Selection criteria prioritised conceptual accessibility, relevance to the target content, the transformation of natural language into mathematical language, and suitability for the time available in each class session.

Instruments

Three instruments were used: a pre-test, a post-test, and a student satisfaction questionnaire. The pre-test was administered before the intervention to assess prior knowledge of systems of linear equations and consisted of 10 scored items. Item responses were coded dichotomously using a checklist-type rule, with 1 assigned to a correct answer and 0 to an incorrect answer, yielding a raw possible score range from 0 to 10.

The post-test was administered after the instructional use of Khan Academy to assess achievement in the same content area. Although it was organised into 10 item prompts, one prompt was divided into two separately scored subparts, resulting in 11 scored components in total. As in the pre-test, item responses were coded dichotomously (1 = correct, 0 = incorrect), yielding a raw possible score range from 0 to 11. Because the pre-test and post-test did not contain exactly the same number of scored components, they were not strictly parallel forms.

The student satisfaction questionnaire was applied after the intervention and consisted of 10 Likert-type items scored on a five-point scale ranging from 1 (Never) to 5 (Always), with a total possible score range from 10 to 50. For transparency, the manuscript distinguishes between the raw item-level coding of the achievement tests and the archived summary scores used in the statistical analyses. Table 1 summarises the structure, response format, scoring rule, and raw score range of the study instruments.

Table 1

Structure, Response Format, Scoring Rule, and Raw Score Range of the Study Instruments

Instrument	Purpose	Scored items / components	Response format	Scoring rule	Raw possible score range
Pre-test	Assessment of prior knowledge of systems of linear equations before the intervention	10	Objective and short constructed-response items	1=correct 0=incorrect	0–10
Post-test	Assessment of achievement in systems of linear equations after the intervention	11 scored components derived from 10 item prompts	Objective and contextualised problem-solving items; one prompt divided into two scored subparts	1=correct 0=incorrect	0–11

Instrument	Purpose	Scored items / components	Response format	Scoring rule	Raw possible score range
Student satisfaction question-naire	Assessment of students' satisfaction with Khan Academy after the intervention	10	Likert-type items	1=Never 2=Rarely 3=Sometimes 4=Almost always 5=Always	10–50

Note. The post-test was organised into 10 item prompts, but one prompt was divided into two separately scored subparts, resulting in 11 scored components.
Source. Own research.

For editorial clarity and to comply with journal formatting standards, the complete list of instrument items (Table A1) is provided as supplementary material in Appendix.

Content Review and Psychometric Caution

The instruments were reviewed by three experts in terms of pertinence, relevance, and clarity. However, the archived material available for the present revision only confirms the experts' overall approval of the instruments and does not preserve the original item-by-item ordinal rating matrix required to compute an inter-rater concordance coefficient such as Kendall's W. For this reason, expert review is reported as qualitative evidence of content appraisal rather than as a formal coefficient of agreement. Likewise, although the statistical reviewer recommended a Rasch analysis, the archived dataset available for the present revision does not include a complete Rasch calibration output. Consequently, the manuscript now reports only the classical test theory evidence that can be verified directly from the original files.

Data Analysis

Data was processed using IBM SPSS Statistics 27. For the revised manuscript, the analysis was reorganised to prioritise cognitively meaningful summary scores rather than fragmented item-type frequencies. Descriptive statistics were calculated for the pre-test, post-test, gain score, and satisfaction score. Normality was assessed through the Kolmogorov-Smirnov test, as reported in the original study. The principal inferential analysis used a paired-samples t-test to compare pre-test and post-test performance. To improve interpretability, the revised results report the mean difference, its 95% confidence interval, and Cohen's dz. Satisfaction with Khan Academy was analysed as a continuous 10-item Likert-type scale and was explored as a secondary variable through correlation with score gain.

Methodological Limitation of Test Comparability

A major methodological limitation concerns test comparability. The pre-test consisted of 10 items, whereas the post-test comprised 11 items because one content component was split into two scored parts in the final assessment. Although both instruments were intended to assess the same content domain, they were not strictly parallel forms. Therefore, the observed pre-post improvement should be interpreted as evidence of performance gain within the instructional context rather than as a comparison between perfectly equivalent measurement forms.

Ethical Considerations

The research process followed the ethical principles stated in the manuscript. Participants were informed about the objectives of the study, the voluntary nature of participation, and the academic use of the data. Digital informed consent was obtained, confidentiality was protected through anonymised records, and results were reported only in aggregate form.

Results

Descriptive Summary of the Main Variables

Table 2 summarises the main descriptive statistics. Before the intervention, cadets obtained a mean pre-test score of 5.85 (SD = 3.03). After the intervention, the mean post-test score increased to 16.30 (SD = 1.57). The mean gain was 10.45 points. Student satisfaction with Khan Academy was high when analysed as a continuous scale, with a mean total score of 47.69 out of 50, equivalent to a mean item score of 4.77 out of 5.

Table 2
Descriptive Statistics for the Main Study Variables

Variable	n	Mean	SD	Median	Min	Max	95% CI
Pre-test score	65	5.85	3.03	6.00	0.00	10.67	5.10-6.60
Post-test score	65	16.30	1.57	16.00	13.33	20.00	15.91-16.69
Gain score (post - pre)	65	10.45	3.08	10.33	4.00	17.33	9.69-11.21
Satisfaction total score (10-50)	65	47.69	3.00	49.00	38.00	50.00	46.95-48.44
Satisfaction means item (1-5)	65	4.77	0.30	4.90	3.80	5.00	4.69-4.84

Source. Own research.

Normality Assessment

The original manuscript reported separate Kolmogorov-Smirnov tests for the pre-test and post-test averages, with non-significant results for both variables (pre-test: $D = 0.107$, $p = .064$; post-test: $D = 0.132$, $p = .070$). These findings supported the use of a parametric repeated-measures comparison.

Pre-Test Versus Post-Test Comparison

A paired-samples t-test showed a statistically significant increase in mathematics performance after the use of Khan Academy (see Table 3). The mean difference between post-test and pre-test scores was 10.45 points (95% CI [9.69, 11.21]). This difference was statistically significant, $t(64) = 27.39$, $p < .001$, with a very large effect size (Cohen's dz = 3.40).

Table 3

Paired-Samples Comparison of Pre-Test and Post-Test Scores

Comparison	Mean difference	SD of difference	95% CI of difference	t	df	p	Cohen's dz
Post-test - Pre-test	10.45	3.08	9.69-11.21	27.39	64	< .001	3.40

Source. Own research.

Learning Levels

To facilitate institutional interpretation, pre-test and post-test scores were also classified according to the achievement bands used by the Ecuadorian Land Force. Before the intervention, all participants were in the insufficient band. After the intervention, only one participant remained in that band, while the distribution shifted towards regular, good, and very good levels (see Table 4).

Table 4

Distribution of Participants Across Institutional Learning Levels

Learning level	Pre-test, n (%)	Post-test, n (%)
Very good	0 (0.0)	12 (18.5)
Good	0 (0.0)	26 (40.0)
Regular	0 (0.0)	26 (40.0)
Insufficient	65 (100.0)	1 (1.5)

Source. Own research.

Student Satisfaction with Khan Academy

Student satisfaction with the platform was high. The mean total score was 47.69 out of 50, indicating a very favourable evaluation of the learning experience. Because the questionnaire was originally structured as a 10-item Likert-type scale, the revised manuscript reports satisfaction as a continuous variable rather than as percentage bands in a pie chart. This change provides a more coherent psychometric interpretation of the instrument.

Exploratory Association Between Satisfaction and Score Gain

An exploratory Pearson correlation was conducted to examine whether students who reported higher satisfaction with Khan Academy also showed greater performance gains. The association between satisfaction total score and gain score was positive but not statistically significant ($r = .197$, $p = .116$). This result suggests that, although overall satisfaction with the platform was high, individual differences in satisfaction did not meaningfully explain variation in learning gains in the present sample.

Discussion

The revised results confirm a substantial increase in cadets' mathematics performance after the instructional use of Khan Academy. The improvement was not only statistically significant but also large in magnitude, which strengthens the empirical argument that the platform was associated with meaningful short-term gains in the learning of systems of linear equations within this military higher education context (Ochogboju & Díez-Palomar, 2025; Rueda-Gómez et al., 2023; Saracostti et al., 2025).

At the same time, the revised manuscript adopts a more cautious interpretative stance than the previous version. The study did not include a control group and therefore cannot isolate the intervention effect from all alternative explanations. In addition, the pre-test and post-test were not strictly parallel forms because they contained different numbers of items. For that reason, the findings support an association between the intervention and improved performance, but they should not be overstated as definitive causal proof under fully controlled experimental conditions. This cautious interpretation is also consistent with previous literature indicating that the effectiveness of Khan Academy depends on pedagogical guidance, contextual conditions, and the way digital resources are integrated into instruction (Abreh et al., 2025; Saracostti et al., 2025; Tosi & Theobald, 2026).

The treatment of satisfaction has also been strengthened. In the previous version, satisfaction was presented through percentage categories and a pie chart, which added little analytical value. In the present revision, satisfaction is treated as a continuous scale, which is more consistent with the instrument's structure. Although the overall level of satisfaction was high, the absence of a significant correlation with score gain indicates that satisfaction should be interpreted as a favourable contextual outcome rather than as a direct statistical explanation of achievement gains. This pattern is compatible with studies showing that digital platforms may enhance motivation, confidence, and engagement without necessarily producing a one-to-one correspondence between positive perceptions and measured achievement gains (Anggraini et al., 2026; Huang et al., 2025; Li & Yin, 2025; Ngu et al., 2025).

The present revision also addresses the statistical review by reducing the overemphasis on fragmented item-type frequencies and by concentrating the evidence on summary scores, confidence intervals, and effect size. This makes the results easier to evaluate and more aligned with standard reporting expectations in quantitative educational research. Nevertheless, further psychometric strengthening remains necessary. Future studies should preserve expert-rating matrices to permit Kendall's W estimation and should examine item functioning through Rasch modelling when the test design and archived outputs support such analysis. More broadly, this need for stronger measurement and better documented digital-learning evidence is consistent with recent discussions on the evaluation of technology-supported learning environments in mathematics and digital education research (Wang & Sun, 2025; Weinhandl et al., 2025).

Conclusions

The study shows that the structured use of Khan Academy was associated with marked improvement in mathematics performance among first-year cadets in ESFORSE. The pre-test/post-test comparison yielded a large increase in scores and a very large effect size, suggesting that the platform functioned as a valuable instructional support resource in this specialised educational context.

The revised manuscript also clarifies that the strength of the findings lies in the observed performance gain, not in an inflated psychometric claim. Reliability evidence for the three instruments was acceptable to high, but the available archive did not allow retrospective estimation of Kendall's W or a full Rasch calibration. Consequently, the article now presents only those analyses that can be verified directly from the original study files.

Overall, the evidence supports the educational usefulness of digital platforms in mathematics teaching in military higher education, while also highlighting the need for more rigorous future designs, parallel assessment forms, and deeper psychometric documentation.

Limitations and Recommendations

The study has several limitations. First, the sample was limited to 65 cadets from a single military institution, which restricts generalisability. Secondly, the quasi-experimental design did not include a control group. Thirdly, the pre-test and post-test were not strictly parallel forms because they contained 10 and 11 items, respectively. Fourthly, the archived evidence available for the present revision did not preserve the original expert-rating matrix needed for Kendall's W, nor did it include a complete Rasch output. Finally, the intervention was limited to one mathematical topic and a relatively short implementation period.

Future research should include control or comparison groups, use strictly parallel achievement tests, preserve complete expert-validation matrices, and incorporate item-level psychometric analyses such as Rasch modelling where appropriate. It would also be advisable to extend the intervention to other mathematical domains and to replicate the study in both military and civilian higher education settings to evaluate external validity more rigorously.

REFERENCES

Abreh, M. K., Arthur, F., Akwetey, F. A., & Nortey, S. A. (2025). Modelling STEM students' intention to learn artificial intelligence (AI) in Ghana: A PLS-SEM and fsQCA approach. *Discover Artificial Intelligence*, *5*(1), Article 223. https://doi.org/10.1007/s44163-025-00466-8

Amalric, M., Roveyaz, P., & Dehaene, S. (2023). Evaluating the impact of short educational videos on the cortical networks for mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, *120*(6), Article e2213430120. https://doi.org/10.1073/pnas.2213430120

Anggraini, A., Triyono, M. B., Sudira, P., & Syahbrudin, J. (2026). Bibliometric analysis: Trends and future directions of entrepreneurship learning models. *Multidisciplinary Reviews*, *9*(1), Article e2026003. https://doi.org/10.31893/multirev.2026003

Borchers, J., & da Cunha, M. S. (2025). School performance and inequality of opportunities in Latin America. *Studies in Educational Evaluation*, *86*, Article 101497. https://doi.org/10.1016/j.stueduc.2025.101497

Egodage, U. K., Wijewickrama, C., Gunawardana, S., Basnayaka, A., & Dissanayake, B. (2025). Use of online resources to study physiology by preclinical medical students: An experience from a developing country. *Advances in Physiology Education*, *49*(3), 742–748. https://doi.org/10.1152/advan.00061.2025

Fruett, F., Pereira Barbosa, F., Cardoso Zampolli Fraga, S., & Ivo Aragao Guimaraes, P. (2024). Empowering STEAM Activities with artificial intelligence and open hardware: The BitDogLab. *IEEE Transactions on Education*, *67*(3), 462–471. https://doi.org/10.1109/TE.2024.3377555

Gleichmann, J., Kubitschke, H., & Schnauß, J. (2025). Impact of a mathematical pre-course on first-year physics students. *European Journal of Science and Mathematics Education*, *13*(3), 172–190. https://doi.org/10.30935/scimath/16363

González-Rogado, A. B., Ramos-Gavilán, A. B., Rodríguez-Esteban, M. A., & Garcia-Holgado, A. (2025). Perception disparity between women and men on the gender gap in STEM at a Spanish university. *Revista Iberoamericana de Tecnologias Del Aprendizaje*, *20*. https://doi.org/10.1109/RITA.2024.3522255

Huang, J., Zhong, Y., & Chen, X. (2025). Adaptive and personalized learning in STEM education using high-performance computing and artificial intelligence. *Journal of Supercomputing*, *81*, Article 981. https://doi.org/10.1007/s11227-025-07481-7

Karimov, A., Nolte, A., Chounta, I.-A., Saarela, M., & Kärkkäinen, T. (2025). Improving learning experience through process re-engineering: Khan Academy localization into Azerbaijani. *Frontiers in Education*, *9*, Article 1419362. https://doi.org/10.3389/feduc.2024.1419362

Lemos, C. C. dos S., Dias, L. O. F., Barbosa, P. S. de A., Marques, E. A. G., Ferraz, R. L., & Nalon, G. H. (2024). Digital tools used on the teaching-learning process in geotechnical engineering. *Soils and Rocks*, *47*(2), Article e2024004923. https://doi.org/10.28927/SR.2024.004923

Li, C., & Yin, W. (2025). Interactive online learning method for students based on artificial intelligence. *Discover Artificial Intelligence*, *5*(1), Article 201. https://doi.org/10.1007/s44163-025-00401-x

Ngu, P. N. H., Nam, N. D., Cuong, L. M., & Thao, T. T. P. (2025). Mathematical modelling in higher education: Evolving research and emerging trends (1980–2023). *Infinity Journal*, *14*(2), 393–418. https://doi.org/10.22460/infinity.v14i2.p393-418

Ochogboju, A. O., & Díez-Palomar, J. (2025). Modeling concrete and virtual manipulatives for mathematics teacher training: A case study in ICT-enhanced pedagogies. *Information*, *16*(8), Article 698. https://doi.org/10.3390/info16080698

Poesia, G., & Goodman, N. D. (2023). Peano: Learning formal mathematical reasoning. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *381*(2251), Article 20220044. https://doi.org/10.1098/rsta.2022.0044

Rueda-Gómez, K. L., Rodríguez-Muñiz, L. J., & Muñiz-Rodríguez, L. (2023). Performance and mathematical self-concept in university students using Khan Academy. *Heliyon*, *9*(4), Article e15441. https://doi.org/10.1016/j.heliyon.2023.e15441

Saracostti, M., Miranda, E., Guzmán, P., Schnettler, B., Concha-Salgado, A., de Toro, X., Lara, L., & Suazo, I. (2025). Exploring school engagement and academic outcomes over time: The role of contextual factors in Chile. *Psychology in the Schools*, *62*(9), 3257–3272. https://doi.org/10.1002/pits.23538

Tosi, M. D. L., & Theobald, M. (2026). Online learning from data streams via decentralized and asynchronous SGD. *Future Generation Computer Systems*, *175*, Article 108052. https://doi.org/10.1016/j.future.2025.108052

Traverso-Condori, L. C., Reyes-Arco, R. E., Reyes-Villalba, E., Torres-Gonzales, S. B., Maraza-Quispe, B., Tinco-Túpac, S. T., & Torres-Loayza, J. L. (2024). Evaluation of the impact of an online educational platform on the development of mathematical competencies in elementary school students. *Revista Iberoamericana de Tecnologias Del Aprendizaje*, *19*, 397–406. https://doi.org/10.1109/RITA.2024.3465021

Wang, G., & Sun, F. (2025). A review of generative AI in digital education: Transforming learning, teaching, and assessment. *International Journal of Information and Communication Technology*, *26*(19), 102–127. https://doi.org/10.1504/IJICT.2025.146701

Weinhandl, R., Andic, B., Baldinger, S., Mayrhofer, J., & Helm, C. (2025). The ICAP model in the context of employing digital technologies for teaching mathematics: An analysis of teachers' strategies to stimulate students' cognitive Engagement moderated by age, gender and experience. *Computers in the Schools*, *0*(0), 1–25. https://doi.org/10.1080/07380569.2025.2544051

Yu, X., & Yang, D. (2024). The influence of mobile technology on STEM education student learning outcomes. *International Journal of Interactive Mobile Technologies (iJIM)*, *18*(20), 37–50. https://doi.org/10.3991/ijim.v18i20.50837

Zaidi, A., Barbareschi, G., Kunze, K., Pai, Y. S., & Yamaoka, J. (2025). TIEboard: A digital educational tool for kids geometric learning. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, *9*(2), Article 62. https://doi.org/10.1145/3729478

Appendix

The appendix is included to ensure transparency and replicability of the measurement instruments without affecting the readability of the main manuscript.

Table A1

Full Item List of the Study Instruments: Original Wording, English Translation, Response Format, and Scoring Criteria

Instrument	Item	Original Spanish wording	English translation	Response format	Scoring / measurement
Pre-test	1	Un sistema de ecuaciones lineales es:	A system of linear equations is:	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Pre-test	2	En general, las ecuaciones fraccionarias se resuelven transformándolas en	In general, fractional equations are solved by transforming them into	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Pre-test	3	Para resolver un sistema de ecuación, Luis hace lo siguiente: despeja "y" de la 2da. ecuación y la expresión obtenida la reemplaza en la 1era. Este método se conoce como:	To solve a system of equations, Luis does the following: he isolates "y" from the second equation and substitutes the resulting expression into the first one. This method is known as:	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Pre-test	4	Cuál es la respuesta del siguiente sistema: 1) $3x + 2y = 7$; 2) $4x - 3y = -2$	What is the solution to the following system: 1) $3x + 2y = 7$; 2) $4x - 3y = -2$?	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Pre-test	5	Por dos latas de horchata y 3 bolsas de papas me han cobrado cinco dólares. ¿Cuál de las siguientes expresiones no puede representar el enunciado?	I was charged five dollars for two cans of horchata and three bags of chips. Which of the following expressions cannot represent this statement?	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect

Instru-ment	Item	Original Spanish wording	English translation	Response format	Scoring / measure-ment
Pre-test	6	¿En qué método de resolución se multiplican las ecuaciones del sistema por números convenientes para que al sumarlas se elimine una de las incógnitas?	In which solution method are the equations in the system multiplied by convenient numbers so that, when added, one of the unknowns is eliminated?	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Pre-test	7	Plantee el sistema de ecuaciones que le permita resolver el problema de la imagen, resuélvalo y seleccione la respuesta correcta: Hemos adquirido 3 aeronaves de ala fija y 2 de ala rotatoria, ambas de origen francés, por $2,5 millones de dólares. Por otro lado, se han adquirido 2 de ala fija y 5 de ala rotatoria, ambas de origen israelí, por $3,4 millones de dólares. ¿Cuál de los siguientes sistemas permite determinar el precio promedio de una aeronave de ala fija y de una de ala rotatoria para las Fuerzas Armadas de Ecuador?	Set up the system of equations that allows you to solve the problem in the image, solve it, and select the correct answer: Three fixed-wing aircraft and two rotary-wing aircraft, both of French origin, were acquired for USD 2.5 million. In addition, two fixed-wing aircraft and five rotary-wing aircraft, both of Israeli origin, were acquired for USD 3.4 million. Which of the following systems determines the average price of one fixed-wing aircraft and one rotary-wing aircraft for the Armed Forces of Ecuador?	Contextualised multiple-choice problem	Dichotomous scoring: 1 = correct, 0 = incorrect

Instru-ment	Item	Original Spanish wording	English translation	Response format	Scoring / measure-ment
Pre-test	8	Teniendo en cuenta el siguiente problema responde cada interrogante: Se ha comprado 3 libras de pólvora y 20 detonadores de acero por $1,45 de la marca TNT y, ayer, 2 libras de pólvora y 50 detonadores de acero por $2. Determinar ¿Cuál es el precio de la libra de pólvora y detonadores? a) ¿Cuáles son las variables del problema? b) ¿Cuáles son los datos conocidos y desconocidos del problema? c) ¿Cómo podría resolver este problema?	Considering the following problem, answer each question: Three pounds of gunpowder and 20 steel detonators were bought for USD 1.45 from the TNT brand and, yesterday, 2 pounds of gunpowder and 50 steel detonators were bought for USD 2. Determine the price of a pound of gunpowder and the detonators. a) What are the variables in the problem? b) What are the known and unknown data in the problem? c) How could you solve this problem?	Short constructed-response item	Dichotomous scoring: 1 = correct, 0 = incorrect
Pre-test	9	Selecciona el enunciado que podría corresponder al sistema:	Select the statement that could correspond to the system:	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect

Instrument	Item	Original Spanish wording	English translation	Response format	Scoring / measurement
Pre-test	10	Resuelve el siguiente problema usando el sistema de ecuaciones en forma gráfica. El entrenamiento de tiro de las tropas militares se realiza con munición real o con polígonos virtuales de tiro... ¿Cuál es el costo de cada disparo real y virtual? Considere que el efecto de cada disparo es similar, tanto en entrenamiento real como simulado.	Solve the following problem using the system of equations graphically. Military troop shooting training is carried out either with real ammunition or through virtual shooting ranges... What is the cost of each real and virtual shot? Assume that the effect of each shot is similar in both real and simulated training.	Graph-based constructed-response item using the provided tables/graph	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	1	Beto es un viajero aventurero. Salta en paracaídas dos veces en cada isla que visita y tres veces en cada península que visita. En la última década, Beto saltó en paracaídas 45 veces en las 19 islas y penínsulas que visitó. ¿Cuántas islas y penínsulas visitó Beto?	Beto is an adventurous traveller. He parachutes twice on each island he visits and three times on each peninsula he visits. In the last decade, Beto parachuted 45 times across the 19 islands and peninsulas he visited. How many islands and peninsulas did Beto visit?	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	2	Ben es 20 años mayor que Daniel. Ben y Daniel se conocieron hace dos años. Hace tres, Ben tenía 3 veces la edad de Daniel.	Ben is 20 years older than Daniel. Ben and Daniel met two years ago. Three years ago, Ben was three times Daniel's age.	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	3	Despeja m y escoge la respuesta correcta: $-7 + 4m + 10 = 15 - 2m$	Solve for m and choose the correct answer: $-7 + 4m + 10 = 15 - 2m$	Objective multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect

Instru-ment	Item	Original Spanish wording	English translation	Response format	Scoring / measure-ment
Post-test	4	Resuelve el sistema de ecuaciones por el método de reducción.	Solve the system of equations using the elimination method.	Objective multiple-choice item with the system shown in the original instrument	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	5	Resuelve el sistema de ecuaciones por el método de susti-tución.	Solve the system of equations using the substitution method.	Objective multiple-choice item with the system shown in the original instrument	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	6	Luis caminó de su casa a la parada del autobús a una velocidad promedio de 5 km/h. Inmedi-atamente se subió al autobús escolar y viajó a una velo-cidad promedio de 60 km/h hasta que llegó a la escuela. La distancia total de su casa a la escuela es 35 kilómetros, y el viaje completo duró 1.5 horas. ¿Cuántos kilómetros recorrió Luis caminando? ¿Cuántos kilómetros recorrió Luis en el autobús?	Luis walked from his home to the bus stop at an average speed of 5 km/h. He then boarded the school bus and travelled at an average speed of 60 km/h until he arrived at school. The total distance from his home to school is 35 kilometres, and the full trip lasted 1.5 hours. How many kilometres did Luis walk? How many kilometres did he travel by bus?	Objective multiple-choice contex-tualised problem	Dichotomous scoring: 1 = correct, 0 = incorrect

Instrument	Item	Original Spanish wording	English translation	Response format	Scoring / measurement
Post-test	7	Courtney caminó desde su casa a la playa a una velocidad constante de 4 kilómetros por hora, y luego de la playa al parque a una velocidad constante de 5 kilómetros por hora. La caminata completa tomó 2 horas y la distancia total que Courtney recorrió fue 8 kilómetros. Sea b el número de horas que le llevó a Courtney caminar de su casa a la playa y sea p el número de horas que le llevó caminar de la playa al parque. ¿Cuál sistema de ecuaciones representa esta situación?	Courtney walked from her house to the beach at a constant speed of 4 kilometres per hour, and then from the beach to the park at a constant speed of 5 kilometres per hour. The full walk took 2 hours and the total distance Courtney travelled was 8 kilometres. Let b be the number of hours it took Courtney to walk from her house to the beach and p the number of hours it took her to walk from the beach to the park. Which system of equations represents this situation?	Contextualised multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	8	Lindsay es 5 años más joven que Mark. Hace siete años, la suma de sus edades era 31. Sea l la edad de Lindsay y sea m la edad de Mark. ¿Cuál sistema de ecuaciones representa esta situación?	Lindsay is 5 years younger than Mark. Seven years ago, the sum of their ages was 31. Let l be Lindsay’s age and m be Mark’s age. Which system of equations represents this situation?	Contextualised multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	9A	¿Cómo podemos obtener el sistema B a partir del sistema A? Escoge una respuesta.	How can we obtain system B from system A? Choose one answer.	Objective multiple-choice subitem	Dichotomous scoring: 1 = correct, 0 = incorrect
Post-test	9B	Con base en la respuesta anterior, ¿son equivalentes los sistemas? En otras palabras, ¿tienen la misma solución?	Based on the previous answer, are the systems equivalent? In other words, do they have the same solution?	Objective multiple-choice subitem	Dichotomous scoring: 1 = correct, 0 = incorrect

Instrument	Item	Original Spanish wording	English translation	Response format	Scoring / measurement
Post-test	10	Resuelve el siguiente sistema de ecuaciones en forma gráfica, estima la solución del sistema de ecuaciones. Puedes usar la gráfica interactiva para encontrar la solución.	Solve the following system of equations graphically and estimate its solution. You may use the interactive graph to find the solution.	Graph-based multiple-choice item	Dichotomous scoring: 1 = correct, 0 = incorrect
Satisfaction questionnaire	11	La nueva estrategia de enseñanza-aprendizaje basada en la plataforma Khan Academy te permitió una mayor comprensión sobre cómo resolver sistema de ecuaciones utilizando los distintos métodos.	The new teaching–learning strategy based on the Khan Academy platform allowed you to gain a better understanding of how to solve systems of equations using different methods.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always
Satisfaction questionnaire	12	El uso del Khan Academy te facilitó para plantear un sistema de ecuaciones que le permita recrear el contexto de los problemas militares de la vida cotidiana.	The use of Khan Academy made it easier for you to formulate a system of equations that recreates the context of everyday military problems.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always
Satisfaction questionnaire	13	Con la participación en el Khan Academy obtuviste información actualizada sobre los conceptos de los sistemas de ecuaciones lineales.	By participating in Khan Academy, you obtained updated information on the concepts of systems of linear equations.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always
Satisfaction questionnaire	14	El material ofrecido en el Khan Academy te facilitó la interpretación del nivel de las diversas representaciones gráficas, tabulares y simbólicas realizadas.	The material provided in Khan Academy made it easier for you to interpret the various graphical, tabular, and symbolic representations used.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always

Instru-ment	Item	Original Spanish wording	English translation	Response format	Scoring / measure-ment
Satisfaction question-naire	15	Con la ayuda del Khan Academy lograste una mejor interpretación del concepto de sistema de ecuaciones lineales a través de la visualización de videos con conteni-dos algebraicos.	With the help of Khan Academy, you achieved a better understanding of the concept of systems of linear equations through viewing videos with algebraic content.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always
Satisfaction question-naire	16	La información publicada en el Khan Academy te permitió mejorar las habilidades a través de las evaluaciones del proceso por cada temática.	The information posted on Khan Academy allowed you to improve your skills through process assessments for each topic.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always
Satisfaction question-naire	17	El Khan Acad-emy como recurso didáctico te permitió el desarrollo de mayores habilidades y capacidades para resolver problemas que involucren el concepto de sistema de ecuaciones lineales.	Khan Academy, as a teaching resource, enabled you to develop greater skills and abilities to solve problems involving the concept of systems of linear equations.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always
Satisfaction question-naire	18	Tu tiempo de estudio mejoró al intervenir en la plataforma Khan Academy.	Your study time improved through participation in the Khan Academy platform.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always

Instrument	Item	Original Spanish wording	English translation	Response format	Scoring / measurement
Satisfaction questionnaire	19	La participación en el Khan Academy te motivó a intercambiar ideas en el aula; a través del proceso comunicativo bidireccional (docente-estudiante) sobre el concepto de sistema de ecuaciones lineales y a su vez a trabajar en equipo.	Participation in Khan Academy motivated you to exchange ideas in the classroom through the bidirectional communication process (teacher–student) about the concept of systems of linear equations and, at the same time, to work as a team.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always
Satisfaction questionnaire	20	Al emplear el Khan Academy como herramienta de didáctica de apoyo al proceso de enseñanza aprendizaje se facilitó la comunicación entre quienes intervinieron en esta experiencia (docente-estudiante).	By using Khan Academy as a didactic support tool in the teaching–learning process, communication between those involved in this experience (teacher–student) was facilitated.	Likert-type item	Likert-type scale: 1 = Never, 2 = Rarely, 3 = Sometimes, 4 = Almost always, 5 = Always

Note. For some pre-test and post-test items, the original instrument includes equations, option sets, tables, or figures embedded visually. In those cases, this table reproduces the full verbal stem and identifies the response format and scoring rule, while the original graphical material remains in the archived instrument file.
Source. Own research.