The Amplifier Effect: Why Guided ChatGPT Predicts Academic Achievement

Julián Andrés Quimbayo Castro
Corporación Universitaria del Huila - CORHUILA
Systems Engineering Program
Neiva – Huila, Colombia
julian.quimbayo@corhuila.edu.co

Eilen Lorena Pérez Montero
Corporación Universitaria del Huila - CORHUILA
Systems Engineering Program
Neiva – Huila, Colombia
eilen.perez@corhuila.edu.co

Jose Miguel Llanos Mosquera
Corporación Universitaria del Huila - CORHUILA
Systems Engineering Program
Neiva – Huila, Colombia
jmllanosm@corhuila.edu.co

Alvaro Hernan Alarcon Lopez
Corporación Universitaria del Huila - CORHUILA
Systems Engineering Program
Neiva – Huila, Colombia
alvaro.alarcon@corhuila.edu.co

Edisney García Perdomo
Corporación Universitaria del Huila - CORHUILA
Systems Engineering Program
Neiva – Huila, Colombia

Abstract

The integration of Large Language Models (LLMs) in higher education presents significant pedagogical challenges. While research has explored various applications, a gap exists in understanding their impact on high-level cognitive tasks like research idea formulation, and how students' pre-existing skills influence outcomes. This study employed a quasi-experimental design with 207 engineering undergraduates, divided into a control group, an unguided ChatGPT group, and a guided ChatGPT group. A binary classification pipeline was developed to predict high-quality research proposals. The optimal model, a Logistic Regression classifier, achieved 93% accuracy, demonstrating high predictive power. The results revealed a critical 'Amplifier Effect.' Feature importance analysis from the model demonstrated that the most significant predictor of achievement was not the AI tool itself, but a student's initial skill, specifically Specific_objective_pretest. The guided use of ChatGPT (Groups_3) acted as a powerful amplifier for this competence, emerging as another top positive predictor, while being in the control group was a strong negative predictor. The findings conclude that the pedagogical framework is more critical than the AI tool itself, suggesting its primary value is to amplify, not supplant, foundational student competence. This study provides empirical evidence for designing structured AI interventions that enhance student skills.

Keywords: ***Large Language Models (LLMs), Artificial Intelligence in Education, Instructional Scaffolding, Predictive Modeling, Amplifier Effect.***

I. INTRODUCTION

The advent of Large Language Models (LLMs) such as ChatGPT represented a transformative force in higher education[1], [2]. While these tools offered significant potential to support complex cognitive tasks, their integration posed pedagogical challenges, including concerns over academic integrity and the potential erosion of critical thinking skills [3], [4]. Consequently, the focus of educational research shifted from whether these tools should be permitted to how they could be leveraged effectively to foster genuine learning and improve academic achievement[5], [6], [7]. A critical area with limited empirical investigation was the application of LLMs to the foundational stage of academic work: the formulation of a research idea.

Prior to this study, existing literature had largely concentrated on the use of AI for discrete tasks like providing

feedback, generating assessment items, or acting as a learning resource [8], [9], [10], which left a gap in understanding how these tools influenced the high-level cognitive process of developing a viable research problem. It was not yet empirically established whether unguided, exploratory use of ChatGPT was as effective as structured, scaffolded interaction. Furthermore, the determinative role of a student's foundational skills in the success of AI-based interventions had not been sufficiently quantified.

This study addressed this question through a comparative analysis of three student groups: a control group without access to ChatGPT, an experimental group with unguided access, and a second experimental group that utilized ChatGPT with specific instructional scaffolding. The research was designed to evaluate the hypothesis that a scaffolded intervention would yield superior outcomes, and to quantify the predictive power of group membership relative to the students' pre-existing abilities.

Through the application of a binary classification model to distinguish between "High" and "Not High" quality research proposals, the analysis revealed a critical insight termed the "Amplifier Effect." The results indicated that the most significant predictor of academic achievement was not the AI tool itself, but the student's initial skill in defining specific research objectives. The study demonstrated that guided use of ChatGPT acted as a powerful amplifier for this pre-existing competence, yielding outcomes that significantly outweighed those from unguided use. This finding suggested that the primary pedagogical value of AI tools lay not in their capacity to replace student skill, but to magnify it when deployed within a structured framework.

The remainder of this paper is organized as follows. Section II reviews related work on artificial intelligence in education and instructional scaffolding. Section III details the experimental methodology, including the dataset, group design, and machine learning pipeline. Section IV presents the classification results and the feature importance analysis. Finally, Section V discusses the implications of the "Amplifier Effect" for AI pedagogy and concludes the paper.

II. LITERATURE REVIEW

A. *The Transformative and Disruptive Role of LLMs in Higher Education*

The emergence of Large Language Models (LLMs) was widely recognized as a pivotal moment for higher education. Scholarly discourse rapidly converged on their potential to revolutionize academic processes [1], [2], prompting a comprehensive rethinking of established educational practices [11]and leading to extensive reviews of the state of the field [12]. The integration of these models was seen not merely as an incremental change but as a strategic transformation with the capacity to reshape teaching, learning, and research paradigms.

Concurrently, this technological integration introduced significant pedagogical and ethical challenges [3]. The discourse was marked by concerns over academic integrity and the potential for misuse [13], framing the integration as both a potential "educational reboot" and a significant disruption [14], [15]. This dual-faceted landscape, characterized by both immense opportunity and considerable risk, underscored the urgent need for empirical research to guide effective and responsible implementation in academic settings.

B. *Current Applications and Student Interactions with LLMs*

In response to this new landscape, research began to explore specific pedagogical applications for LLMs. These investigations focused on leveraging the models as interactive tutors for problem-solving [5], tools for generating competence-based assessments and questions [8], [16], platforms for delivering effective feedback [10], and engines for creating playful, game-based learning environments [6]. These studies demonstrated the functional versatility of LLMs in performing structured educational tasks.

Parallel to these application-focused studies, another line of inquiry investigated student interactions with these tools. Research focused on the factors influencing student acceptance and use [17], their decision-making processes when choosing between LLMs and traditional search methods [9], and their overall perspective as learners navigating this new technological landscape [4]. This work highlighted that the effectiveness of any AI tool was intrinsically linked to how students perceived, accepted, and ultimately engaged with it.

C. *The Research Gap: From Prompt Engineering to Predicting Achievement*

Despite this progress, a significant portion of the literature concentrated on the operational aspects of LLM integration, such as developing effective prompts [18], designing new curricula [18], or exploring administrative applications [19]. This focus left a discernible gap in understanding the impact of LLMs on high-level, ill-defined cognitive processes, such as the formulation of a novel research idea [20], [21]. It was not yet clear how the unstructured nature of creative academic tasks aligned with the capabilities of these models.

Perhaps the most critical gap, however, was the under-examination of the student's own pre-existing skills as a predictive variable. While predictive modeling had been used to analyze academic achievement based on sociodemographic or pedagogical data [22], [23], [24], this approach had not been systematically applied to measure the relative importance of student skill versus a specific AI intervention. The foundational learning competencies of students [7] remained an under-quantified factor in the success of AI-supported tasks.

Therefore, this review identified a crucial gap in the literature. While research had explored the applications and challenges of LLMs, there was a need for empirical studies that (a) compared the effectiveness of guided versus unguided LLM use on high-level cognitive tasks, and (b) quantified the predictive importance of students' initial skills relative to the intervention itself. To address this gap, this study investigated the following research question: **RQ1.** *Was the guided use of*

ChatGPT a more significant predictor of academic achievement than its unguided use, particularly when accounting for a student's initial skills?

III. METHODOLOGY

This section details the research methodology, including the description of the participant cohort, the data collection instruments, the experimental design, and the data analysis pipeline employed to answer the research question.

A. Dataset and Participants

A cohort of 207 undergraduate students from the Faculty of Engineering at Corporación Universitaria del Huila - CORHUILA participated in a mandatory Research Methodology course for the study. The final dataset comprised the complete records of all participants who completed both stages of the intervention.

B. Experimental Design and Instruments

A quasi-experimental, pre-test/post-test design was employed to compare the effectiveness of different pedagogical interventions. The cohort of 207 participants was distributed equitably across three groups: a Control Group (no AI), an Experimental Group 1 (unguided AI use), and an Experimental Group 2 (guided AI use). The primary instrument for both assessments was a detailed analytic rubric designed to evaluate the quality of research idea formulation. The rubric employed a scale from 0.0 to 5.0 for each criterion.

C. Data Analysis and Predictive Modeling

A machine learning pipeline was created to analyze data and identify academic achievement predictors. The process was structured as follows:

1) ***Target Variable Definition:*** A target variable, Quality, was created from the Posttest scores to enable binary classification. A score of 3.5 or higher was categorized as 'High' (1), while scores below this threshold were categorized as 'Not High' (0). This threshold clearly separates high-quality work based on the data distribution.

2) ***Feature Engineering:*** The original feature set derived from the pre-test rubric items was expanded through feature engineering. This process involved creating aggregate features (e.g., Pretest_Average, Pretest_Sum) and interaction features (e.g., Intro_x_Context_pretest) to capture more complex relationships within the data.

3) ***Modeling Pipeline Construction:*** A machine learning pipeline was constructed to ensure robust and reproducible data processing. The pipeline integrated three key stages: (1) a preprocessing step using a ColumnTransformer to apply MinMaxScaler to numerical features and OneHotEncoder to the categorical Grupos feature; (2) a resampling step using RandomOverSampler to address class imbalance in the training data; and (3) the classification model itself.

4) ***Model Training and Hyperparameter Tuning:*** A suite of seven distinct classification algorithms was evaluated: Logistic Regression, K-Nearest Neighbors, Support Vector Machine, Decision Tree, Random Forest, Gradient Boosting, and XGBoost. The optimal hyperparameters for each model were determined using GridSearchCV, implemented with a RepeatedStratifiedKFold cross-validation strategy (10 splits, 3 repeats) to ensure model stability. The f1_weighted score was employed as the primary metric for selecting the best-performing model pipeline.

5) ***Model Evaluation and Interpretation:*** The performance of the final, optimized model was evaluated using a comprehensive classification report (precision, recall, F1-score) and a confusion matrix. Finally, to answer the primary research question, feature importances were extracted from the best-performing model. This analysis allowed for the quantification of the predictive power of each variable, including the experimental group assignments, thereby revealing the key determinants of academic achievement in this context.

IV. RESULTS

1) ***Target Variable Definition:***

For binary classification analysis, continuous post-test scores were transformed into a categorical target variable (Quality). This dichotomization distinguished between achievement levels using a data-informed threshold of 3.5. Observations were classified as 'High' (≥3.5) or 'Not High' (<3.5) and subsequently encoded as 1 and 0, respectively, for machine learning compatibility. This binary target variable facilitated all subsequent modeling.

Figure 1 illustrated this classification boundary and revealed distinct performance patterns across experimental conditions. The Control group exhibited a distribution centered below the threshold, with most participants scoring in the 'Not High' range (2.5-3.5), establishing the baseline performance. Experimental Group 1 (Unguided AI) demonstrated a significant rightward shift, with substantially more participants exceeding the 3.5 threshold compared to Controls. Experimental Group 2 (Guided AI) showed the most pronounced effect, with a strongly right-skewed distribution and minimal 'Not High' classifications.

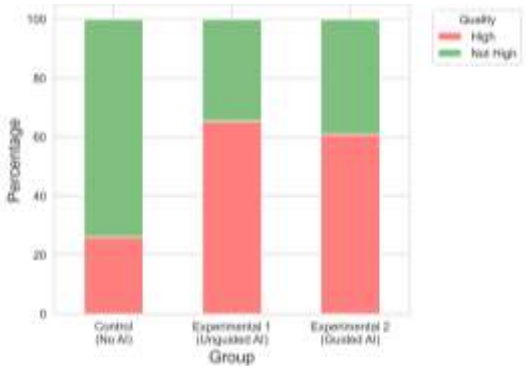
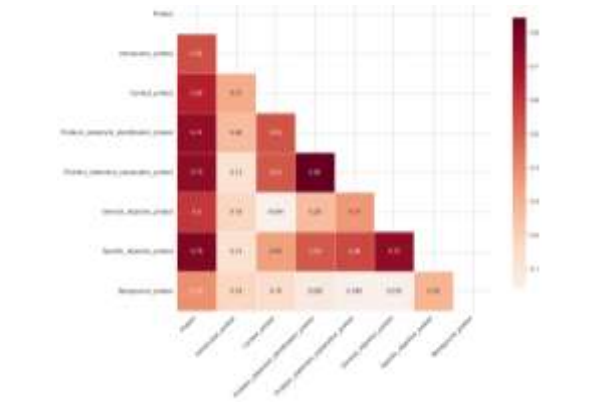


Fig. 1 Quality Distribution by Group

Quantitative analysis confirmed these observations. The Control group established a robust baseline with 75% of participants in the 'Not High' category. Experimental Group 1 showed significant improvement, with 'Not High' classifications decreasing to approximately 40%, allowing 60% to exceed the threshold. Experimental Group 2 exhibited the most dramatic improvement, with fewer than 10% classified as 'Not High' and over 90% achieving 'High' status.

Figure 2 presented a comprehensive analysis of pre-test and post-test performance across groups. The box plots in the upper left panel revealed comparable pre-test scores across all conditions, confirming the effectiveness of randomization. Post-test scores, however, showed a clear stepwise



improvement pattern from Control to Unguided AI to Guided AI groups. The upper right panel quantified this improvement, with median score gains of approximately 0.9, 1.2, and 1.2 points for the respective groups. The lower left violin plot further illustrated the distinct post-test score distributions, with progressively higher medians and upper quartiles across the three conditions.

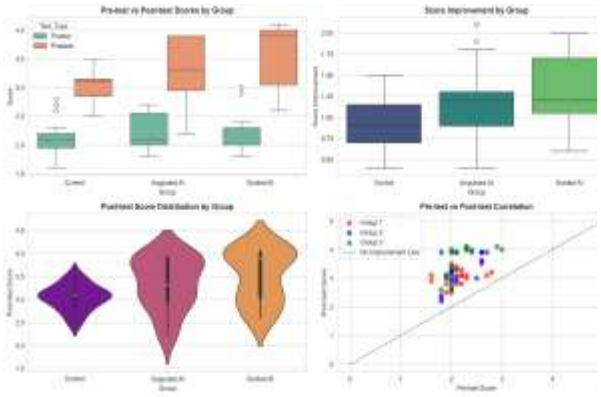


Fig. 2 Pre-test and Post-test Complete Analysis

The scatter plot (lower right panel) demonstrated that participants across all groups achieved post-test scores above the no-improvement line, with Guided AI participants (Group 3) consistently showing the highest gains. This visualization confirmed that pre-test scores were not predictive of group assignment, eliminating baseline knowledge as a confounding variable.

The correlation analysis (Figure 3) revealed strong positive correlations between Group assignment and outcome measures (r=0.78 with Post-Test; r=0.71 with Quality), statistically confirming group assignment as a powerful performance predictor. The near-perfect correlation between Post-Test and Quality (r=0.89) validated the variable transformation. Importantly, the negligible correlation between Pre-Test and Group (r=-0.03) confirmed well-balanced groups with comparable baseline knowledge. These analyses collectively demonstrated the substantial and varied impact of AI-based interventions on student achievement, confirming group assignment as a critical predictive variable and validating the binary Quality variable's sensitivity to intervention-induced performance variations.

Fig. 3 Correlation Matrix of Pre-test components

*2) **Feature Engineering:***

Following correlation analysis, feature engineering was implemented to enhance model predictive capacity. Three aggregate features (Pretest_Average, Pretest_Sum, Pretest_Std) were created to capture overall performance metrics and distribution characteristics of baseline knowledge.

Four interaction features were developed based on conceptual relationships in research writing: Intro_x_Context_pretest leveraged the observed correlation (r=0.31) between introductory and contextual elements; Problem_x_Objective_pretest combined problem identification with goal formulation; Background_x_Intro_pretest integrated background knowledge with introductory writing; and Specific_x_General_Obj_pretest captured the hierarchical relationship between objective types (r=0.75).

This engineering approach expanded the feature space to 16 predictors, incorporating domain-specific knowledge about research writing structure while preserving the experimental design integrity. The enhanced feature set provided classification algorithms with more nuanced data for distinguishing between performance outcomes across the experimental conditions.

*3) **Modeling Pipeline Construction:***

A machine learning pipeline was implemented with preprocessing (MinMaxScaler for numerical features, OneHotEncoder for 'Groups'), RandomOverSampler for class balancing, and model training components.

*4) **Model Training and Hyperparameter Tuning:***

Seven classification algorithms (Logistic Regression, KNN, SVM, Decision Tree, Random Forest, Gradient Boosting, XGBoost) were evaluated using three train-test partitioning strategies (60-40, 70-30, 80-20 splits) with stratification preserved. Hyperparameter optimization utilized GridSearchCV with 3 repetitions × 10 splits and f1_weighted as the optimization metric, ensuring balanced performance across classes while mitigating cross-validation variation effects.

*5) **Model Evaluation and Interpretation:***

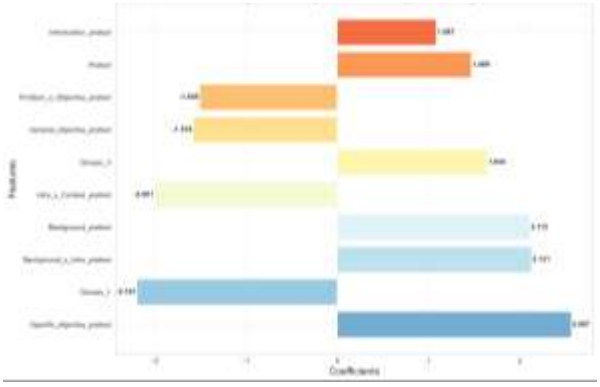


Table 1 presents the classification performance metrics across different train-test split ratios. A clear pattern of improved predictive accuracy emerged as the proportion of training data increased, demonstrating the models' enhanced learning capability with larger training sets.

TABLE I. PERFORMANCE METRICS ACROSS DIFFERENT TRAIN-TEST SPLIT RATIOS

Split Ratio	*Best Model*	*Accuracy*	*F1Score (Class 0)*	*F1Score (Class 1)*
60 - 40	SVC	0.71	0.73	0.69
70 - 30	SVC	0.90	0.90	0.91
80 - 20	***Logistic Regression***	***0.93***	***0.92***	***0.93***

The 60-40 split yielded moderate performance with SVC achieving 71% accuracy. Class-specific metrics revealed a slight imbalance in prediction quality, with better recall for 'Not High' outcomes (0.79) than for 'High' outcomes (0.64). This suggested the model was more effective at identifying students who would not achieve high performance than those who would excel.

A substantial performance improvement occurred with the 70-30 split, where SVC achieved 90% accuracy with remarkably balanced precision and recall across both classes (approximately 0.90). This balance indicated the model's equal effectiveness in identifying both high-achieving and lower-performing students.

The 80-20 split produced the highest overall accuracy (93%) with Logistic Regression emerging as the optimal model. This configuration demonstrated perfect precision (1.00) for 'Not High' predictions and perfect recall (1.00) for 'High' predictions. These results indicated that with sufficient training data, the model could identify all students who would achieve high performance (Class 1) while maintaining high precision in identifying those who would not.

Figure 4 illustrates the top 10 most influential features based on the coefficients from the optimal Logistic Regression model. The analysis revealed that both baseline skills and experimental group assignment were critical predictors of academic achievement.

Fig. 4 Top 10 Most Important Features

Specific_objective_pretest emerged as the most influential predictor with the highest positive coefficient (2.567), indicating that students' initial ability to formulate specific research objectives strongly predicted high achievement. This was closely followed by Background_x_Intro_pretest (2.131) and Background_pretest (2.113), highlighting the importance of contextual knowledge and its integration with introductory writing skills.

Notably, Groups_3 (Guided AI intervention) showed a substantial positive coefficient (1.644), confirming that participation in the guided ChatGPT group was a powerful predictor of high achievement. Conversely, Groups_1 (Control) exhibited a strong negative coefficient (-2.191), indicating that absence of AI assistance significantly predicted lower performance outcomes.

Among the engineered features, Intro_x_Context_pretest showed a negative coefficient (-2.001), suggesting a complex interaction effect where the relationship between introductory and contextual elements required careful balance. Similarly, Problem_x_Objective_pretest (-1.505) demonstrated that the interaction between problem identification and objective setting had nuanced effects on performance prediction.

These feature importance findings quantitatively answered the research question: the guided use of ChatGPT (Group 3) was indeed a more significant predictor of academic achievement than its unguided use (Group 2), particularly when accounting for students' initial skills. The models successfully distinguished between the three experimental conditions, with the guided intervention consistently associated with the highest probability of achieving 'High' performance classification.

V. DISCUSSIONS AND CONCLUSIONS

This study was designed to answer a critical question in the era of generative AI: Was the guided use of ChatGPT a more significant predictor of academic achievement than its unguided use, particularly when accounting for a student's initial skills? The results of our predictive modeling provide a clear, albeit nuanced, answer. The findings confirmed that the guided intervention was indeed a more powerful predictor of success than unguided use, but not in isolation. The central finding of this research is the identification of what we term the "Amplifier Effect": AI tools like ChatGPT function most effectively not as a replacement for foundational academic skills, but as a powerful magnifier of them.

The feature importance analysis (Fig. 4) provided compelling evidence for this effect. The single most influential predictor of achieving a 'High' quality research proposal was not the AI intervention itself, but a pre-existing

student skill: Specific_objective_pretest. This indicates that a student's foundational ability to formulate precise research objectives was the primary determinant of success. The experimental group assignments functioned as modulators of this core competence. Participation in the guided AI group (Groups_3) was a strong positive predictor, demonstrating that the structured scaffolding enabled students to leverage their initial skills to achieve superior outcomes. Conversely, being in the control group (Groups_1) was the strongest negative predictor, confirming that the absence of any AI tool significantly limited achievement on this complex task.

These results suggest a change in basic assumptions for AI pedagogy. The discourse should move beyond simply "prompt engineering" towards "structured pedagogical design." The unguided use of ChatGPT (Group 2) yielded improvements over the control group but was markedly less effective than the guided intervention. This implies that merely providing access to powerful AI tools is insufficient. The true value is unlocked when educators design structured, scaffolded experiences that guide students to use these tools as a means to amplify their developing cognitive abilities. The negative coefficients for some interaction features, such as Intro_x_Context_pretest, further suggest that the interplay between skills is complex and that effective scaffolding must help students navigate these nuanced relationships.

While this study provides robust findings, several limitations must be acknowledged. The research was conducted with undergraduate engineering students at a single institution, which may limit the generalizability of the results to other disciplines or educational contexts. Furthermore, the study captured a short-term intervention; longitudinal research is needed to determine the long-term effects of these interventions on skill retention and development. Finally, the rapid evolution of LLMs means that future versions may interact with student skills in different ways.

This study empirically demonstrated that the pedagogical framework surrounding an AI tool is more critical than the tool itself. In response to RQ1, we conclude that the guided use of ChatGPT was a significantly more effective predictor of academic achievement than unguided use. However, its primary function was to amplify students' pre-existing skills—most notably, their ability to define specific research objectives. The "Amplifier Effect" posits that AI's greatest educational potential is realized when it is used to enhance, not supplant, foundational human intellect.

Future research should aim to replicate these findings across diverse academic disciplines and investigate the long-term impacts of scaffolded AI use on student learning. Further exploration into diverse types of scaffolding for various high-level cognitive tasks would also represent a valuable contribution to the field, helping educators to integrate AI strategically and effectively into their curricula.

REFERENCES

[1] M. Diab Idris, X. Feng, y V. Dyo, «Revolutionizing Higher Education: Unleashing the Potential of Large Language Models for Strategic Transformation», *IEEE Access*, vol. 12, pp. 67738-67757, 2024, doi: 10.1109/ACCESS.2024.3400164.

[2] W. Gan, Z. Qi, J. Wu, y J. C.-W. Lin, «Large Language Models in Education: Vision and Opportunities», en *2023 IEEE International Conference on Big Data (BigData)*, dic. 2023, pp. 4776-4785. doi: 10.1109/BigData59044.2023.10386291.

[3] P. A. P. Dungca, «The Incorporation of Large Language Models (LLMs) in the Field of Education: Ethical Possibilities, Threats, and Opportunities», en *Philosophy of Artificial Intelligence and Its Place in Society*, IGI Global Scientific Publishing, 2023, pp. 78-97. doi: 10.4018/978-1-6684-9591-9.ch005.

[4] K. Zdravkova, F. Dalipi, y F. Ahlgren, «Integration of Large Language Models into Higher Education: A Perspective from Learners», en *2023 International Symposium on Computers in Education (SIIE)*, nov. 2023, pp. 1-6. doi: 10.1109/SIIE59826.2023.10423681.

[5] N. P. Bakas, M. Papadaki, E. Vagianou, I. Christou, y S. A. Chatzichristofis, «Integrating LLMs in Higher Education, Through Interactive Problem Solving and Tutoring: Algorithmic Approach and Use Cases», en *Information Systems*, M. Papadaki, M. Themistocleous, K. Al Marri, y M. Al Zarouni, Eds., Cham: Springer Nature Switzerland, 2024, pp. 291-307. doi: 10.1007/978-3-031-56478-9_21.

[6] S. E. Huber, K. Kiili, S. Nebel, R. M. Ryan, M. Sailer, y M. Ninaus, «Leveraging the Potential of Large Language Models in Education Through Playful and Game-Based Learning», *Educ. Psychol. Rev.*, vol. 36, n.º 1, p. 25, mar. 2024, doi: 10.1007/s10648-024-09868-z.

[7] Y. Assefa, B. T. Moges, y S. A. Tilwani, «Lifelong learning measurement scale (LLMS): development and validation in the context of higher education institutions», *J. Appl. Res. High. Educ.*, vol. 16, n.º 3, pp. 693-705, jul. 2023, doi: 10.1108/JARHE-04-2023-0164.

[8] R. Meissner *et al.*, «LLM-generated competence-based e-assessment items for higher education mathematics: methodology and evaluation», *Front. Educ.*, vol. 9, p. 1427502, oct. 2024, doi: 10.3389/feduc.2024.1427502.

[9] R. R. Divekar, S. Guerra, L. Gonzalez, y N. Boos, «Choosing Between an LLM versus Search for Learning: A HigherEd Student Perspective», 19 de septiembre de 2024, *arXiv*: arXiv:2409.13051. doi: 10.48550/arXiv.2409.13051.

[10] T. Ruwe y E. Mayweg-Paus, «Embracing LLM Feedback: the role of feedback providers and provider information for feedback effectiveness», *Front. Educ.*, vol. 9, p. 1461362, oct. 2024, doi: 10.3389/feduc.2024.1461362.

[11] G. Chhabra, N. Mehdian, y P. Vasishta, «Rethinking Higher Educational Practices in the Age of Artificial Intelligence», en *2024 IEEE 5th India Council International Subsections Conference (INDISCON)*, ago. 2024, pp. 1-6. doi: 10.1109/INDISCON62179.2024.10744297.

[12] H. Crompton y D. Burke, «Artificial intelligence in higher education: the state of the field», *Int. J. Educ. Technol. High. Educ.*, vol. 20, n.º 1, p. 22, abr. 2023, doi: 10.1186/s41239-023-00392-8.

[13] T. Rasul *et al.*, «The role of ChatGPT in higher education: Benefits, challenges, and future research directions», *J. Appl. Learn. Teach.*, vol. 6, n.º 1, Art. n.º 1, may 2023, doi: 10.37074/jalt.2023.6.1.29.

[14] O. S. Ogunleye, «ChatGPT Implications on Higher Education: Educational Apocalypse or Educational Reboot? A Developing Countries Perspective», en *2023 International Conference on Computational Science and Computational Intelligence (CSCI)*, dic. 2023, pp. 1685-1690. doi: 10.1109/CSCI62032.2023.00278.

[15] M. Neumann, M. Rauschenberger, y E.-M. Schön, «"We Need To Talk About ChatGPT": The Future of AI and Higher Education», en *2023 IEEE/ACM 5th International Workshop on Software Engineering Education for the Next Generation (SEENG)*, may 2023, pp. 29-32. doi: 10.1109/SEENG59157.2023.00010.

[16] S. T. Vu, H. T. Truong, O. T. Do, T. A. Le, y T. T. Mai, «A ChatGPT-based approach for questions generation in higher education», en *Proceedings of the 1st ACM Workshop on AI-Powered Q&A Systems for Multimedia*, en AIQAM '24. New York, NY, USA: Association for Computing Machinery, jun. 2024, pp. 13-18. doi: 10.1145/3643479.3662056.

[17] A. Habibi, M. Muhaimin, B. K. Danibao, Y. G. Wibowo, S. Wahyuni, y A. Octavia, «ChatGPT in higher education learning: Acceptance and use», *Comput. Educ. Artif. Intell.*, vol. 5, p. 100190, 2023, doi: 10.1016/j.caeai.2023.100190.

[18] M. Nazari y G. Saadi, «Developing effective prompts to improve communication with ChatGPT: a formula for higher education

stakeholders», *Discov. Educ.*, vol. 3, n.º 1, p. 45, abr. 2024, doi: 10.1007/s44217-024-00122-w.
[19] Y. Li, «The Potential Application of ChatGPT in Higher Education Management», *Lect. Notes Educ. Psychol. Public Media*, vol. 25, n.º 1, pp. 200-208, nov. 2023, doi: 10.54254/2753-7048/25/20230750.
[20] E. Kasneci *et al.*, «ChatGPT for good? On opportunities and challenges of large language models for education», *Learn. Individ. Differ.*, vol. 103, p. 102274, abr. 2023, doi: 10.1016/j.lindif.2023.102274.
[21] F. E. Oguz, M. N. Ekersular, K. M. Sunnetci, y A. Alkan, «Can Chat GPT be Utilized in Scientific and Undergraduate Studies?», *Ann. Biomed. Eng.*, vol. 52, n.º 5, pp. 1128-1130, may 2024, doi: 10.1007/s10439-023-03333-8.
[22] E.-L. Pérez-Montero y J.-A. Quimbayo-Castro, «Predicción del logro académico en clases espejo: análisis sociodemográfico y pedagógico con minería de datos», *Rev. Científica*, vol. 49, n.º 1, pp. 79-98, feb. 2024, doi: 10.14483/23448350.21820.
[23] A. H. A. López, J. M. L. Mosquera, J. A. Q. Castro, E. G. Perdomo, I. K. A. Ramirez, y D. H. S. Rincón, «LMS Effect on Programming Learning: A Data Mining Case Study», en *2024 IEEE VII Congreso Internacional en Inteligencia Ambiental, Ingeniería de Software y Salud Electrónica y Móvil (AmITIC)*, sep. 2024, pp. 1-8. doi: 10.1109/AmITIC62658.2024.10747596.
[24] J. A. Q. Castro, E. G. Perdomo, Á. H. A. Lopez, D. L. Rodríguez, y Y. G. A. C. Laiton, «Exploring Older Adults' Acceptance and Use of Online Transactions: A Machine Learning-Based Analysis», en *2023 VI Congreso Internacional en Inteligencia Ambiental, Ingeniería de Software y Salud Electrónica y Móvil (AmITIC)*, oct. 2023, pp. 1-8. doi: 10.1109/AmITIC60194.2023.10366368.