applied sciences

MDPI

Article

Analyzing Higher Education Students' Prompting Techniques and Their Impact on ChatGPT's Performance: An Exploratory Study in Spanish

José Luis Carrasco-Sáez [1], Carolina Contreras-Saavedra [2], Sheny San-Martín-Quiroga [1], Carla E. Contreras-Saavedra [3] and Rhoddy Viveros-Muñoz [1,*]

[1] Departamento de Electrónica e Informática, Universidad Técnica Federico Santa María, Concepción 4030000, Chile; jose.carrascosa@usm.cl (J.L.C.-S.); shsanmartin@doctoradoia.cl (S.S.-M.-Q.)
[2] Facultad de Educación, Universidad Católica de la Santísima Concepción, Campus San Andrés, Concepción 4070409, Chile; carolina.contreras@ucsc.cl
[3] Facultad de Ciencias de la Rehabilitación y Calidad de Vida, Universidad San Sebastián, Concepción 4081339, Chile; carla.contreras@uss.cl
* Correspondence: rhoddy.viveros@usm.cl

Abstract

Generative artificial intelligence is reshaping how people interact with digital technologies, emphasizing the need to develop effective skills for engaging with it. In this context, prompt engineering has emerged as a critical skill for optimizing AI-generated outputs. However, research on how higher education students interact with these technologies remains limited, particularly in non-English-speaking contexts. This exploratory study examines how 102 higher education students in Chile formulated prompts in Spanish and how their techniques influenced the responses generated by ChatGPT (free version 3.5). A quantitative analysis was conducted to assess the relationship between prompt techniques and response quality. Two emergent prompt engineering strategies were identified: the Guide Contextualization Strategy and the Specific Purpose Strategy. The Guide Contextualization Strategy focused on providing explicit contextual information to guide ChatGPT's responses, aligning with few-shot prompting, while the Specific Purpose Strategy emphasized defining the request's purpose, aligning with structured objective formulation strategies. The regression analysis indicated that the Guide Contextualization Strategy had a greater impact on response quality, reinforcing the importance of contextual information in effective interactions with large language models. As an exploratory study, these findings provide preliminary evidence on prompt engineering strategies in Spanish, a relatively unexplored area in artificial intelligence education research. Based on these results, a methodological framework is proposed, encompassing four key dimensions: grammatical skills; prompt strategies; response from the large language model; and evaluation of response quality. This framework lays the groundwork for future artificial intelligence digital literacy interventions, fostering critical and effective engagement with generative artificial intelligence while also highlighting the need for further research to validate and expand these initial insights.

Keywords: prompt engineering; natural language processing; AI in higher education; generative AI; Spanish grammar performance; generative artificial intelligence; AI literacy

Academic Editors: Venkat N. Gudivada and Rajvardhan Patil

Received: 28 May 2025
Revised: 2 July 2025
Accepted: 4 July 2025
Published: 8 July 2025

Citation: Carrasco-Sáez, J.L.; Contreras-Saavedra, C.; San-Martín-Quiroga, S.; Contreras-Saavedra, C.E.; Viveros-Muñoz, R. Analyzing Higher Education Students' Prompting Techniques and Their Impact on ChatGPT's Performance: An Exploratory Study in Spanish. *Appl. Sci.* **2025**, *15*, 7651. https://doi.org/10.3390/app15147651

1. Introduction

The rise in generative artificial intelligence (GenAI) marks one of the most transformative technological shifts in recent decades, reshaping how individuals learn, communicate, and solve problems. Its rapid proliferation underscores the need to reassess the skills required for citizens to interact effectively with these advanced tools. This new scenario began when the company OpenAI (San Francisco, CA, USA) launched ChatGPT (version 3.5) in 2022. However, other tech giants such as Google (Mountain View, CA, USA), Microsoft (Redmond, WA, USA), and Meta (Menlo Park, CA, USA) already have their own GenAI platforms, widely accessible to people around the world.

The rapid development of artificial intelligence (AI) has brought notable changes to various fields, particularly education. Within this context, ChatGPT has gained prominence due to its advanced conversational capabilities and its potential to enhance teaching and learning processes. As a chatbot powered by AI, ChatGPT can offer tailored and interactive assistance to students by addressing complex queries, providing feedback on academic tasks, and facilitating open-ended dialog. Thanks to its flexibility and practical applications, this tool holds the potential to reshape traditional models of higher education [1].

These advancements are referred to as large language models (LLMs), which are trained on vast volumes of text to process and generate natural language. They are considered part of the field of natural language processing (NLP), which enables models to 'understand,' generate, and analyze human language. This represents a significant advantage, as it allows users to interact with LLMs using natural language rather than code, removing the need for programming skills to effectively engage with them.

Typically, the process of interacting with a GenAI involves a user providing information in various formats, such as text or audio, through a prompt. The GenAI then processes the input using a large language model and generates an output, also in multiple formats. The user does not know how the GenAI analyzes their input to transform and generate the output. Therefore, the only way to interact with the GenAI is by using the prompt.

Various authors [2,3] point out that managing the use of prompts is a skill that can be acquired and improved by following certain rules to enhance the richness of grammatical information provided in the input, incorporating context, roles, output formats, and other aspects. In this context, concepts such as prompt engineering, prompt strategies, and prompt techniques have been emerging. However, distinguishing between strategies and techniques is not straightforward, as they sometimes share the same names or are grouped into a generic category.

On the other hand, considering that the use of prompts in the context of GenAI is relatively new, it is not yet clear which strategies or techniques yield the best results in the quality of responses provided by GenAI. Educational communities are beginning to explore how these technologies can be integrated. Nonetheless, the lack of time and resources makes it difficult to develop a road map that facilitates their systematic integration and measures the real impact of GenAI on the teaching-learning process. Students are also facing the challenge of interacting with GenAI, requiring them to develop new skills, such as effective prompt design. These skills are not always intuitive and require strong guidance and support to leverage the use of GenAI in a critical and ethical manner.

In summary, the arrival of GenAI is redefining the concept of digital literacy, shifting from basic technology skills to the ability to communicate effectively with complex artificial intelligence systems. This shift implies a transformation in educational programs, emphasizing proper interaction with these tools while addressing risks such as technological dependence and the lack of critical thinking skills.

Although several studies have analyzed and classified prompt techniques [2–7], little is still known about the specific strategies employed by students to leverage the opportunities

brought by GenAI. While the potential of GenAI in education is widely acknowledged, further research is needed to understand how students interact with these tools, particularly in Spanish, given that most LLMs are primarily trained in English.

Understanding how higher education students interact with GenAI is a critical first step in developing strategies to support their effective use of these tools. This exploratory study examines the prompting techniques employed by the students while working with the free version of ChatGPT (version 3.5). By analyzing both circumstantial and request-based techniques, the research identifies emergent dimensions of prompt strategies and provides a comprehensive characterization of the simple and complex strategies utilized in prompt construction. These techniques are particularly relevant as they influence the quality and effectiveness of interactions with GenAI, particularly in Spanish—a language that poses distinct grammatical and contextual challenges for LLMs. Furthermore, the study explores the relationship between the types of techniques employed and the quality of the responses generated. Beyond assessing whether students fully understand the techniques they apply, the research examines how the interaction and interrelation of these techniques, grouped into emergent strategies, impact the outputs produced by GenAI. By offering a deeper understanding of students' prompting behaviors, the findings contribute to a foundational framework for designing educational interventions that promote critical, ethical, and effective engagement with LLMs in non-English-speaking contexts.

Despite the growing integration of generative AI tools like ChatGPT in academic settings, there remains limited understanding of how students interact with these systems—particularly in non-English contexts. Most of the research on prompt engineering and AI literacy has focused on English-speaking users, often overlooking the linguistic and cognitive challenges faced by Spanish-speaking students. Moreover, there is a lack of pedagogical frameworks to guide students in crafting effective prompts and critically evaluating AI responses. This study was motivated by the need to explore these gaps and contribute empirical evidence to support the development of culturally and linguistically appropriate AI literacy programs in higher education.

To achieve this purpose, the study was guided by the following research questions:

(1) What prompting strategies do higher education students use when interacting with ChatGPT in Spanish? (2) How do the types of techniques used in prompt formulation relate to the perceived quality of the AI-generated responses? (3) What patterns of simple and complex prompting strategies emerge from students' interactions, and how might these inform future AI literacy initiatives?

2. Related Work

The emergence of GenAI has sparked a wave of innovation across various sectors, transforming how people and organizations coexist with technology. This accelerated development is challenging organizations to understand how these tools could bring competitive advantages. However, the implications of the massive adoption of GenAI are still underexplored, especially in the education sector, where the challenges are even more complex due to its structured nature and occasionally rigid curricula.

In this context, the pathways for integrating these new technologies may depend on the level of digital maturity within educational communities, particularly regarding their capacity to incorporate and innovate with digital tools in academic, cultural, and knowledge management processes, ultimately generating a significant impact on their external environment. Consequently, implementation proposals may range from comprehensive initiatives that encompass the entire community to smaller, informal practices based on trial and error. An important difference between artificial intelligence (AI) before GenAI and GenAI itself lies in the fact that, with the latter, we do not necessarily know where the

data used to train and operate its models comes from. Therefore, the most critical way to interact with GenAI is through the prompt.

Nonetheless, the integration of ChatGPT into higher education is influenced by a range of elements that determine its acceptance and actual usage among both students and faculty members [4–7]. Gaining insight into these elements is crucial for harnessing the full potential of ChatGPT and facilitating its effective incorporation into academic settings. Empirical research has explored multiple dimensions related to ChatGPT's adoption, such as its technological features, users' attitudes and experiences, institutional frameworks, and the broader educational landscape.

Considering the above, in the last two years, various investigations have been conducted to build a knowledge base that enables understanding and analyzing the interactions between users and prompts. Within this emerging field, concepts such as prompts, prompt strategies, prompt techniques, and prompt engineering have emerged.

In the domain of LLMs, a prompt refers to an input provided by a user to a GenAI model to guide and trigger a specific output. Typically, the prompt consists of a piece of text used to communicate with LLMs. It can take the form of a question, a statement, a set of instructions, or any other input that serves as the basis for the model's response [8–11].

A prompt also functions as a translational bridge between human communication and the computational abilities of LLMs. It consists of natural language instructions that operate as an intermediary language, translating human intentions into machine-readable tasks. Beyond basic directives, prompts establish a contextual framework for the interaction, signaling the relevance of certain elements and delineating the structure and content expected in the model's output [12].

Understanding the elements of a prompt and their importance is crucial for effective interactions with GenAI. These elements include the following, according to Giray [13]: (1) a task or specific instruction that guides the model's behavior toward the desired results; (2) external contextual information that provides the model with additional knowledge, helping it generate more accurate and relevant outputs; (3) the input data or question that the user wants the model to process and respond to; and (4) specifications regarding the type or format of the desired output. These elements serve as a foundation for designing effective prompts that maximize the capabilities of GenAI models.

Building upon this foundational understanding of prompt components, it becomes essential to explore how these elements can be intentionally combined and optimized through the practice of prompt engineering.

Prompt engineering is the strategic process of designing inputs (prompts) that guide LLMs to generate specific, relevant outputs. It is akin to providing the right questions or commands to ensure that the desired answers are obtained. Prompt engineers leverage the predictive capabilities and extensive training data of AI models and combine them with the nuances of human communication to strategically influence model behavior without the need for task-specific training [14,15]. Effective prompt engineering tactically employs the various best practices of prompt design as needed to maximize the performance of LLMs, making them more accurate and relevant for specific tasks. It involves not just the creation of prompts but also their testing and refinement based on feedback and desired outcomes.

Similarly, Eager and Brunton [16] provided guidance on producing instructional text to enhance the development of high-quality outputs from LLMs in higher education. To facilitate the process of prompt engineering, they recommend including six key components in written prompts: (a) verb, (b) focus, (c) context, (d) focus and condition, (e) alignment, and (f) constraints and limitations. These categories offer a structured lens for designing effective prompts in educational settings and align with broader efforts to foster AI literacy.

As LLMs are increasingly used for more complex tasks, the process of prompt engineering tends to involve assembling multiple prompts in creative and layered ways. Due to the models' sensitivity to natural language input, slight adjustments in wording, sentence construction, or contextual cues can significantly alter the resulting outputs [17,18].

A prompt technique is a framework that describes how to construct a single prompt, multiple prompts, or a dynamic sequence of interconnected prompts. It can incorporate features such as conditional or branching logic, parallelism, or other architectural considerations to structure a cohesive group of prompts. These techniques may be implemented either manually or through automation, depending on the complexity of the task and the desired level of precision [19].

Among these techniques, two of the most significant are the following: (a) few-shot learning, which involves instructing an LLMs to learn a new task using only a few examples. This approach facilitates spontaneous task delegation and improves model performance [20]. (b) Zero-shot learning, which prompts LLMs to perform tasks without any prior examples. This technique can be further enhanced by fine-tuning the instructions and incorporating reinforcement learning with human feedback [21,22].

Prompt engineering is an iterative process that involves designing, refining, and optimizing input prompts to effectively convey the user's intent to a language model like ChatGPT. This practice often includes modifying or adjusting the prompting technique to improve the quality of the responses. Proper prompt engineering is essential for obtaining accurate, relevant, and coherent outputs, making it a critical skill for users seeking to maximize the potential of GenAI.

As language models continue to advance, mastering prompt engineering has become increasingly important across a wide range of applications [3,23]. Haugsbaken and Hagelia [2] suggest that prompts can vary in their levels of complexity, identifying four distinct levels: (a) basic prompting, (b) prompt reviewing, (c) prompt planning, and (d) prompt scaffolding. (a) Basic prompting: This practice is associated with educational promptization and focuses primarily on input interactions with AI language models. It reflects a student's ability to engage with chatbots in a basic manner, using a 'programming approach' to generate a variety of content or responses, ranging from simple to complex. (b) Prompt reviewing: This level involves an emphasis on the output aspect of input-output engagements. It entails critically examining and refining the outputs provided by AI models to better align them with the intended objectives of the interaction. (c) Prompt planning: At this level, prompting evolves into a more complex and recursive process. Students construct a series of interconnected prompts designed to structure a learning process. This approach moves beyond generic communication with AI models, focusing on creating tailored and purposeful interactions that support specific learning goals.

Various studies are being conducted to understand and classify prompt techniques, as well as to evaluate their effectiveness. Schulhoff et al. [3], using the PRISMA methodology for a systematic review, proposed an ontological taxonomy of text-based prompting techniques. This taxonomy includes 58 techniques distributed across six major categories: (a) zero-shot, (b) few-shot, (c) thought-generation, (d) ensembling, (e) self-criticism, and (f) decomposition.

Giray [13] proposes five prompting techniques: (a) instructive prompt, (b) system prompt, (c) question–answer prompt, (d) contextual prompt, and (e) mixed prompt. An instructive prompt allows the user to employ a guide to direct their writing toward a specific task. A system prompt provides a starting point or context to help users build their content effectively. A question–answer prompt facilitates the structuring of inputs around specific research questions. A contextual prompt offers additional background or details within the prompt to help users focus on particular aspects of their response. Finally, a

mixed prompt combines multiple elements, enabling users to construct their inputs in a comprehensive and understandable manner.

Sawalha, Taj, and Shoufan [24] conducted a study to investigate how undergraduate students used ChatGPT for problem-solving, the prompting strategies they developed, the relationship between these strategies and the model's response accuracy, the presence of individual prompting tendencies, and the influence of gender in this context. The study involved 56 senior students enrolled in the course. The findings revealed that students predominantly employed three types of prompting strategies: single copy-and-paste prompting (SCP), single reformulated prompting (SRP), and multiple-question prompting (MQP). ChatGPT's response accuracy using SRP and MQP was significantly higher than with SCP, with effect sizes of -0.94 and -0.69, respectively.

Recent advancements in prompt engineering have led to the development of diverse techniques aimed at optimizing interactions with conversational LLMs. Fagbohun and Harrison [25] proposed a comprehensive and interdisciplinary framework that categorizes these prompting techniques into seven conceptual groups: logical and sequential processing, contextual understanding and memory, specificity and targeting, meta-cognition and self-reflection, directional and feedback, multimodal and cross-disciplinary, and creative and generative. Each category encompasses a set of strategies tailored to specific interaction goals, such as reasoning through complex problems, eliciting reflective responses, or generating creative content. This taxonomy offers a roadmap for practitioners and educators to select the most appropriate prompting strategies based on their pedagogical objectives or domain-specific applications.

According to Sawalha et al. [24], although there is a growing body of research on GenAI prompts, most studies have focused on their use by researchers rather than by students. Few authors have explored how students interact with GenAI tools such as ChatGPT, highlighting a significant gap in understanding student-specific behaviors and strategies in this context.

In this regard, Knoth et al. [22] examined how AI literacy among higher education students influences their interactions with LLMs. For this study, the author defined AI literacy as "a set of competencies that enables individuals to critically evaluate AI technologies, communicate and collaborate effectively with AI, and use AI as a tool online, at home, and in the workplace" [26] (p. 2). The study involved a sample of 45 students aged between 19 and 35 years. The findings revealed that higher-quality prompt engineering skills significantly predict the quality of LLM outputs, highlighting prompt engineering as an essential skill for the goal-directed use of generative AI tools. Furthermore, the study showed that certain aspects of AI literacy play a crucial role in improving prompt engineering quality and enabling targeted adaptation of LLMs within educational contexts.

Building upon the findings from related works, which highlight the existence of a wide range of techniques for prompt construction, this study proposes a synthesized framework comprising two groups of techniques: circumstantial and request-based. Circumstantial techniques are defined as those that provide context and structure to the prompt, forming a foundational layer to guide the interaction. For this study, 10 circumstantial techniques were identified and defined (see Table 1). On the other hand, request-based techniques refer to those explicitly used to articulate the query or task to the LLMs. The 11 request-based techniques proposed in this framework are detailed in Table 2. This categorization serves as a practical synthesis of the existing body of knowledge, offering a structured approach that can be applied to analyze the work produced by students, as further elaborated in the methodology and results sections.

Table 1. Identified circumstantial techniques.

N°	Circumstantial Technique (CT)	Definition
1	Context	Provides a framework to guide the response. The context can be more or less explicit depending on the amount of additional information to be included. Example: "I'm preparing a pitch for investors. . .".
2	Ambiguity Reduction	Crucial for tasks that may have multiple interpretations. This technique removes unnecessary ambiguities or clarifies potentially confusing terms. Example: "I'm preparing a pitch for investors in a startup focused on renewable energy".
3	Response Format Instructions	Specifies how the response should be structured (e.g., paragraphs, lists, tables). Ensures that the output aligns with the desired style or level of clarity. Example: "Respond with a bulleted list of the key benefits of machine learning."
4	Conditions or Assumptions	Incorporates specific assumptions or conditions that the model must consider to respond appropriately, reducing undesired interpretations. Example: "Assume the reader has basic knowledge of statistics."
5	Goal or Purpose Indicator	Guides the response by specifying the underlying objective or reason behind the prompt. This helps the model align its output with the intended purpose of the request. Example: "I'm planning a workshop for beginners on artificial intelligence and want to make the session engaging. Provide practical and interactive activities I can include." In this case, the objective is to design an engaging and interactive workshop.
6	Style, Tone, or Linguistic Constraints	Defines the style or tone of the response, such as academic, conversational, technical, or humorous. Adjusts the language to suit the intended audience. Example: "Use simple language without technical jargon".
7	Exhaustiveness Indicator	Specifies whether a comprehensive response covering all possibilities is expected or a concise one focusing on key aspects. Example: "Provide an exhaustive list of the main machine learning algorithms".
8	Time or Temporal Perspective	Indicates a specific timeframe or perspective to guide the response (e.g., "currently" or "in the next five years"). Example: "Describe the current state of machine learning in the healthcare industry".
9	Role Indicator	Defines the role from which the response should be constructed, such as "you are a doctor" or "you are a teacher."
10	Perspective or Point of View Indicator	Specifies the perspective for constructing the response, distinct from the role assigned to the model. This may include viewpoints like "analyze from an economic perspective" or "from the public health standpoint." Example: "Explain the impact of AI from the perspective of an ethics specialist".

Table 2. Identified request-based techniques.

N°	Request-Based Technique (RBT)	Definition
11	Direct Request	Involves asking for a specific action in a simple sentence without adding context. Example: "Define what machine learning is".
12	Mixed Request or Multiple Tasks	Combines multiple instructions in a single prompt, such as "Describe machine learning and provide examples of its use in medicine." Enables richer responses.
13	Request with Specified Detail Level or Focus	Specifies the desired level of detail, such as "Briefly describe" or "Explain each step of the learning process in detail." Example: "Provide information related to i, ii, iii".

Table 2. *Cont.*

N°	Request-Based Technique (RBT)	Definition
14	Open-Ended Question Request	Uses open-ended questions like "What does it imply...?" to elicit broad and exploratory responses.
15	Sequential Request	Indicates that the response should address several aspects in a specific sequence, such as "First define, then provide examples, and finally analyze the limitations".
16	Request for Analysis from Multiple Disciplines or Perspectives	Asks the model to approach the response from multiple disciplines or perspectives, valuable for multidisciplinary topics. Example: "Analyze the impact of artificial intelligence from medical, ethical, and technological perspectives".
17	Request for Fact-Based or Data-Specific Response	Seeks a response grounded in concrete facts or known statistics, even hypothetically, for a more objective and evidence-based focus. Example: "Based on recent studies, explain the impact of AI on employment".
18	Request with Extension or Depth Limitation	Specifies a limit on the response length, such as "in less than 100 words" or "in three sentences." Depth can also be limited, like "provide an overview" or "conduct a detailed analysis." Example: "Summarize the applications of neural networks in medicine in less than 100 words".
19	Request with Purpose	Asks for a specific action while indicating the purpose behind the request, such as "Explain xxx so I can prepare for my class".
20	Closed-Ended Question Request	Uses closed-ended questions like "Is machine learning effective in medicine?" to elicit concise answers.
21	Request with Context	Adds context to guide the response, which can range from minimal to extensive depending on the additional information provided. Example: "Create a reading list for a graduate student specializing in computational linguistics".

Building upon the findings from related works, we developed a 21-item checklist through a three-step process grounded in a functional classification of prompt engineering techniques. First, an exploratory review of academic and technical literature was conducted to identify common prompting techniques relevant to educational contexts. From this review, 21 techniques were consolidated based on frequency, conceptual clarity, and applicability in higher education.

In the second step, a functional analysis grouped these techniques into two overarching categories based on their communicative role: (i) circumstantial techniques (n = 10), which provide structural or contextual cues for shaping the model's response; and (ii) request-based techniques (n = 11), which explicitly articulate the action or task expected from the LLM. Each technique was clearly defined and illustrated with an example (see Tables 1 and 2), forming a practical and pedagogically oriented framework for analyzing prompt construction.

This classification informed the design of the 21-item binary checklist used in the study, enabling a systematic analysis of student-generated prompts. The application of this framework is further elaborated in the methodology and results sections.

While this study centers on the functional structuring of prompts, the impact of grammatical and syntactic features on LLM outputs has been explored in detail in a recent publication by Viveros-Muñoz et al. [27], which specifically examines these linguistic aspects.

To better situate this study within existing educational paradigms, it is important to consider how current theories of AI literacy and digital competence frame user interaction with generative AI. Ng [28] defines AI literacy as a multidimensional construct that includes not only technical knowledge but also cognitive and ethical understanding, allowing users to make informed decisions about the use of AI. This concept aligns with recent calls to

develop AI literacy programs that empower students to question, interpret, and strategically engage with AI systems in educational settings.

In parallel, the Digital Competence Framework for Citizens (DigComp 2.2), developed by the European Commission, outlines five competence areas. Of particular relevance are "Digital Content Creation", which emphasizes the ability to produce and adapt digital content critically and responsibly, and "Problem Solving", which highlights the importance of identifying needs and responding through appropriate digital tools [29]. These frameworks offer a conceptual foundation for understanding the competencies required for meaningful engagement with AI-based technologies in education.

Although this study does not directly assess these frameworks, it builds on their principles by exploring how students formulate and refine prompts, interpret AI responses, and evaluate their usefulness—processes that implicitly involve digital content creation, problem-solving, and critical thinking.

3. Materials and Methods

The material used to collect the data was a questionnaire, and the analysis was based on a descriptive exploratory study methodology. The details will be presented in the following points.

3.1. Data Collection

The research sample consisted of 102 higher education students from the Biobío Region in Chile. A convenience sampling approach was employed, with participants recruited based on the accessibility and willingness of the collaborating educational institutions [30]. Students came from two undergraduate programs—rehabilitation (approximately 30%) and computer science (approximately 70%)—at two universities. Although the analysis was conducted on the overall sample, the inclusion of students from both a health-related and a technological program helps to contextualize the educational background of participants.

Although the sample is non-probabilistic, it is considered appropriate for the purposes of this exploratory study, as it includes students with varying levels of familiarity with the use of GenAI. This sampling strategy offers a broad perspective on the prompting strategies employed by students, despite the limitations inherent to its non-random nature.

3.2. Procedure

The data collection procedure followed a structured approach. First, students were welcomed, and the context and objectives of the study were explained. Participants were then presented with an informed consent form, which they reviewed before confirming their voluntary participation. The questionnaire was administered via Google Forms to ensure efficient and anonymous data collection. Students completed the questionnaire using their personal devices—such as mobile phones or tablets—within the classroom environment to maintain consistent technological conditions. The questionnaire comprised three main sections:

- Demographics: This section gathered information such as age, gender, educational institution, academic program, year of study, and prior experience with ChatGPT, providing context for participants' responses.
- Case Analysis: Participants were presented with a common scenario: preparing for a job interview in their field of study. They were instructed to use ChatGPT to generate examples of typical interview questions and tips for effectively answering them.
- Satisfaction Assessment: This section measured participants' satisfaction with the responses generated by ChatGPT.

For the prompt interaction phase, students were asked to write a prompt in the questionnaire, input it into the free version of ChatGPT (3.5), and copy the AI-generated response back into the form. They were then required to evaluate their satisfaction with the response. If satisfied, the test concluded; if dissatisfied, they could rewrite the prompt up to seven iterations until they achieved satisfaction.

Ethical considerations were central to the procedure. Participation was entirely voluntary, and students could withdraw at any time without facing any consequences. Anonymity and confidentiality were rigorously ensured, with all data collected being used exclusively for academic purposes. The entire activity lasted approximately 35 min, providing a structured and controlled environment for data collection and analysis.

After data collection, the responses were analyzed to identify the presence of the 21 proposed techniques (circumstantial and request-based) within the prompts. Each prompt was independently evaluated by two experts in artificial intelligence, who assigned a binary code of 1 if a technique was present or 0 if it was absent (for more details see Section 3.3). The final results were consolidated and systematically tabulated for further analysis.

3.3. Study Design and Analytical Approach

To validate the evaluation instrument, two experts independently analyzed each prompt, assigning a binary code of 1 if a particular technique was identified in the prompt and 0 if it was not. Cases where the two reviewers disagreed were flagged for further analysis. In such instances, a third expert acted as an arbitrator to resolve the discrepancies. To prevent bias, the arbitrator was blinded to the decisions of the initial reviewers, ensuring an impartial assessment of whether the prompt employed the technique in question. All reviewers were specialists in artificial intelligence applied to education and healthcare, providing the necessary expertise to accurately evaluate the use of the proposed techniques.

This study adopts a quantitative approach, structured into descriptive and correlational phases, to address its objectives comprehensively. In the first phase, the study focuses on Objective 1, which seeks to identify emerging dimensions of the prompt strategies employed by higher education students. To achieve this, an Exploratory Factor Analysis (EFA) was conducted, grouping the techniques used into distinct dimensions based on their patterns of co-occurrence.

The second phase addresses Objectives 2 and 3. In relation to Objective 2, the study explores the relationship between the types of techniques students used to formulate prompts and the quality of the responses generated by the AI. This phase includes a descriptive analysis, followed by correlation and regression analyses aimed at identifying patterns and potential causal links. For Objective 3, the focus is on characterizing the simple and complex strategies observed in students' prompt-writing processes. A Rasch model is applied to assess the difficulty of the responses generated by the AI, offering insights into how specific strategies may influence the overall complexity of the interaction.

4. Results

4.1. Emerging Dimensions of Prompt Strategies

To address Objective 1, a tetrachoric correlation analysis was conducted using the statistical software Jamovi (version 2.6.44) to examine the relationships between the dichotomous variables indicating the presence or absence of each technique. Items 3, 6, 7, 10, and 16 were excluded from the analysis due to insufficient response variability. Notably, negative correlations were identified for items 11, 12, 14, and 20, as indicated by the red lines in the network diagram presented in Figure 1.

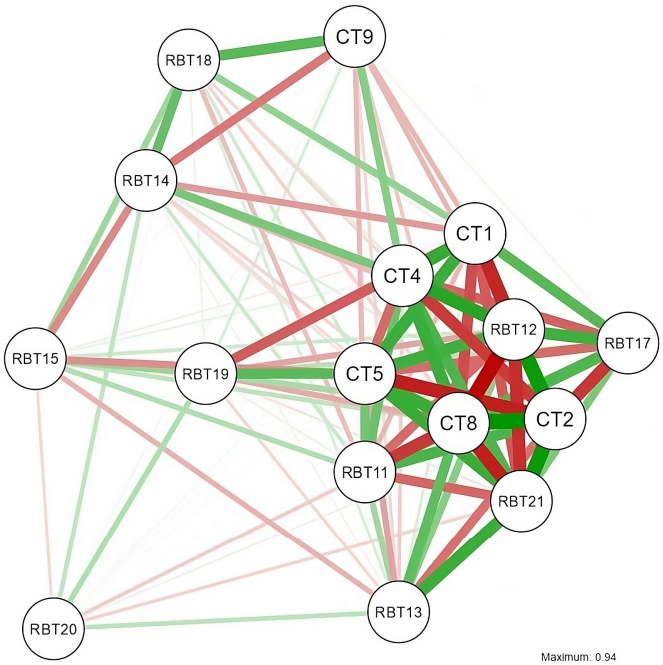

Figure 1. Network structure. The thickness of the lines represents the strength of the correlation between two factors. Green and red lines indicate positive and negative correlations, respectively, among the 16 items (see numbering in Tables 1 and 2). Note: CT: Circumstantial Technique; RBT: Request-Based Technique.

A moderate negative tetrachoric correlation (rt = –0.518) was observed between the use of the direct request technique and the mixed or multi-task request technique. This indicates that as the presence of one of these strategies increases, the presence of the other tends to decrease. In other words, students who frequently used direct prompts tended not to employ complex or multi-tasking formulations simultaneously, suggesting a possible distinction in their approach to interacting with the model.

High positive tetrachoric correlations (>0.7, in Table 3) were identified between indicators 1 and 2, indicators 1 and 21, indicators 2 and 21, and indicators 13 and 21. Conversely, indicator 14 exhibited a strong negative correlation with other techniques.

The RBT14 shows a negative correlation with the other techniques. This inverse association indicates that when an open-ended question is formulated (e.g., "What does it imply...?"), it is less likely that the prompt will include elements intended to frame, constrain, or structure the response. In other words, the use of open-ended questions tends to be accompanied by fewer contextual or structural cues, suggesting a preference for unstructured interaction styles when employing exploratory formulations.

Figure 2 illustrates the process preceding the exploratory factor analysis (EFA), where 69 relevant responses were selected from the 102 participants. Different models were tested: first, considering only positively correlated items, and then including both positive and negative correlations. The EFA was conducted to determine the most suitable model for grouping the techniques (see Figure 2). Model 5 demonstrated a good fit, with an RMSEA of 0.00 (90% CI: [0.00, 0.018]), an excellent TLI of 0.977, a chi-square (χ^2) value of 8.76 (df = 9, p = 0.460), and an explained variance of 45.2%. Meanwhile, Model 3 exhibited the best overall fit, with an RMSEA of 0.00 (90% CI: [0.00, 0.019]), a χ^2 value of 6.05 (df = 1, p = 0.83), and an explained variance of 53.6%. The KMO values were 0.819 for Model 5 and 0.741 for Model 3, indicating strong sampling adequacy. Between the two best-fitting models, Model 3 was ultimately selected due to its stronger theoretical alignment and better conceptual coherence regarding the two emerging factors.

Table 3. Tetrachoric correlations among techniques.

	CT1	CT2	CT4	CT5	CT8	CT9	RBT11	RBT12	RBT13	RBT14	RBT15	RBT17	RBT18	RBT19	RBT20	RBT21
CT1	1.000															
CT2	0.832	1.000														
CT4	0.380	0.123	1.000													
CT5	0.649	0.591	0.132	1.000												
CT8	0.094	−0.081	0.545	0.400	1.000											
CT9	0.424	0.455	0.615	0.077	0.377	1.000										
RBT11	0.148	0.351	−0.036	0.325	0.098	0.091	1.000									
RBT12	0.567	0.393	0.400	0.392	0.277	0.482	−0.518	1.000								
RBT13	0.199	0.037	0.539	−0.012	0.538	0.621	0.016	0.227	1.000							
RBT14	−0.770	−0.715	−0.251	−0.576	−0.125	−0.360	−0.503	−0.289	−0.176	1.000						
RBT15	0.252	0.043	0.504	0.286	0.553	0.299	0.347	−0.129	0.731	−0.235	1.000					
RBT17	0.105	0.153	0.289	0.130	0.656	0.429	0.139	0.003	0.539	−0.105	0.564	1.000				
RBT18	0.453	0.484	0.315	0.379	0.391	0.495	−0.139	0.640	0.610	−0.336	0.311	0.431	1.000			
RBT19	0.450	0.444	0.049	0.617	0.435	−0.054	0.565	−0.123	0.098	−0.468	0.403	0.487	0.193	1.000		
RBT20	−0.032	−0.237	−0.156	−0.025	0.159	−0.079	−0.128	0.056	0.104	−0.305	0.047	0.175	−0.070	−0.042	1.000	
RBT21	0.702	0.783	0.374	0.409	−0.011	0.489	0.381	0.172	0.251	−0.540	0.447	0.300	0.469	0.327	−0.356	1

CT: Circumstantial Technique; RBT: Request-Based Techniques.

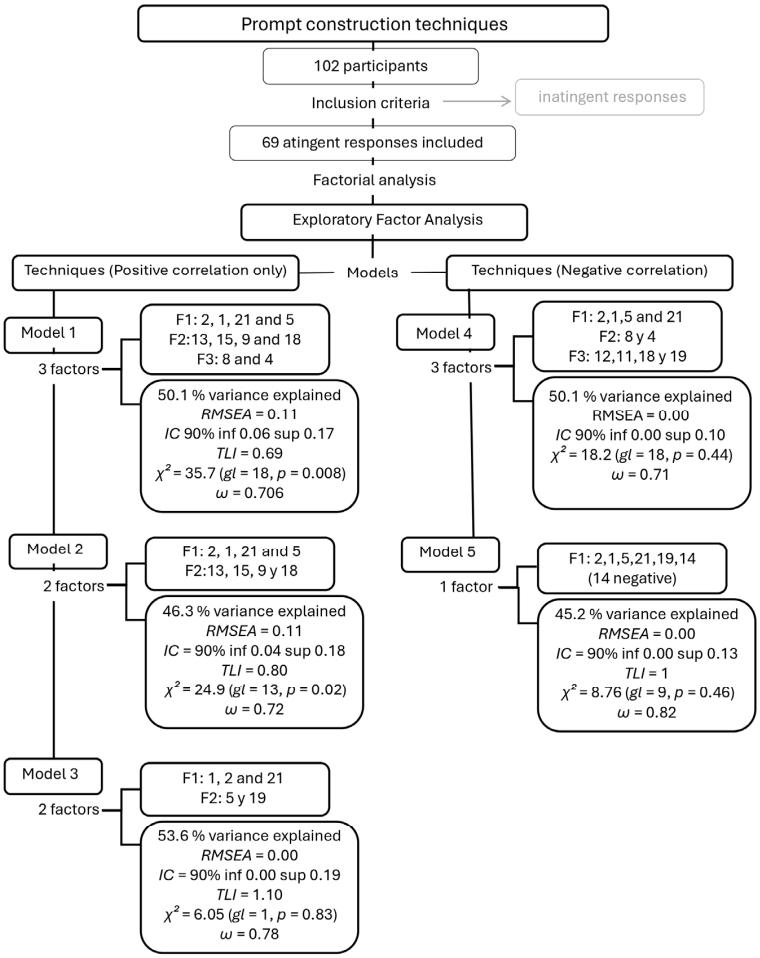

Figure 2. Exploratory analysis flow. This figure shows the sequence of decisions and model tests conducted during the exploratory factor analysis. It highlights the selection of 69 valid responses from 102 participants, the division of techniques by correlation polarity, and the fit statistics for each model tested.

Model 3 revealed two theoretical strategies: the contextual strategy, comprising techniques 1, 2, and 21, and the purpose-related strategy, comprising techniques 5 and 19 (see Table 4).

Table 4. Factor loadings comparison—EFA Model 5 and Model 3.

Technique	Model 5 KMO 0.819		Model 3 KMO 0.741		
	Factor 1	Uniqueness	Factor 1	Factor 2	Uniqueness
CT2 Ambiguity Reduction	0.882	0.223	0.913		0.140
CT1 Context	0.833	0.306	0.748		0.374
RBT21 Request with context	0.622	0.613	0.669		0.591
RBT14 Open-Ended Question Request	−0.727	0.472			
CT5 Goal or Purpose Indicator	0.461	0.787		0.622	0.547
RBT19 Request with Purpose	0.338	0.886		0.606	0.667

Note: The 'maximum likelihood' extraction method was used in combination with a 'promax' rotation.

4.2. Relationships Between Techniques and Quality of Responses

4.2.1. Tetrachoric Correlations

To address Objective 2, the study examined the relationship between techniques used by students and the quality of responses obtained from ChatGPT. Table 5 shows that the quality of correct responses was highly correlated with CT1 (rt = 0.632) and CT2 (rt = 0.656), and moderately correlated with RBT21 (rt = 0.497). This indicates that CT1 and CT2 strongly align with correct responses.

Table 5. Correlation matrix—tetrachoric.

	CT1	CT2	CT5	RBT14	RBT19	RBT21	Val. Response	GPT Use	Strategy Use
CT1	1	0.933	0.680	−0.842	0.550	0.801	0.632	0.538	0.546
CT2	0.933	1	0.609	−0.866	0.510	0.844	0.656	0.440	0.632
CT5	0.680	0.69	1	−0.659	0.632	0.410	0.369	0.142	0.476
RBT14	−0.842	−0.866	−0.659	1	−0.655	−0.676	−0.280	−0.339	−0.442
RBT19	0.550	0.510	0.632	−0.655	1.000	0.299	0.005	0.015	0.108
RBT21	0.801	0.844	0.41	−0.676	0.299	1	0.497	0.275	0.651
Valorisation of the Response	0.632	0.656	0.369	−0.280	0.005	0.497	1.000	0.411	0.301
GPT use	0.538	0.440	0.142	−0.339	0.015	0.275	0.411	1.000	0.583
Strategy use	0.546	0.632	0.476	−0.442	0.108	0.651	0.301	0.584	1

Students' prior use of ChatGPT was moderately correlated with CT1 (rt = 0.538), CT2 (rt = 0.440), and their self-reported use of strategies during the exercise (rt = 0.584). CT1 showed positive associations with CT2 (rt = 0.933), CT5 (rt = 0.680), RBT19 (rt = 0.550), and RBT21 (rt = 0.801). Conversely, RBT14 exhibited strong negative correlations with CT1 (rt = −0.842) and CT2 (rt = −0.866), suggesting that the presence of RBT14 reduced the likelihood of CT1 and CT2 being employed.

4.2.2. Linear Regression Analysis

A linear regression analysis was conducted using a six-point scale (1 = very incorrect to 6 = very correct) to predict response quality based on the strategies employed (see Table 6). Techniques CT1, CT2, and RBT21 were significantly associated with the variable quality of response, $F(1, 67) = 17.7$, $p < 0.001$, with an R^2 value of 0.209, indicating that the model accounted for 20.9% of the variance in the dependent variable. The intercept of the model was 4.98, with an average increase of 1.283 attributable to these strategies. The overall

model was statistically significant, with a standardized coefficient β = 0.457, reflecting a high-moderate effect size [31].

Table 6. Model results—Strategy 1.

Model Fit								
R^2	Adj. R^2	Df	df (res)	F			p	
0.209	0.197	1	67	17.7			<0.001	
ANOVA Omnibus tests								
	SS	Df	F	p			η^2 p	
Model	19.847	1	17.698	<0.001			0.209	
dimensión 2, 1, 21	19.847	1	17.698	<0.001			0.209	
Residuals	75.138	67						
Total	94.986	68						
Parameter Estimates (Coefficients)								
95% Confidence Intervals								
Names	Estimate	SE	Lower	Upper	B	Df	T	p
(Intercept)	4.986	0.127	4.731	5.24	0	67	39.106	<0.001
dimensión 2, 1, 21	1.283	0.305	0.675	1.892	0.457	67	4.207	<0.001

As shown in Table 7, techniques CT5 and RBT19 also exhibited a statistically significant relationship with quality of response, F (1, 67) = 8.58, $p < 0.005$, with an R^2 value of 0.11, accounting for 11% of the variance in the outcome variable. The model yielded an intercept of 4.98, with an average increase of 1.243 associated with these strategies. The effect size was moderate, β = 0.337.

Table 7. Model results—Strategy 2.

Model Fit								
R^2	Adj. R^2	Df	df (res)	F			p	
0.114	0.1	1	67	8.58			0.005	
ANOVA Omnibus tests								
	SS	Df	F	p			η^2 p	
Model	10.788	1	8.585	0.005			0.114	
dimension 5, 19	10.788	1	8.585	0.005			0.114	
Residuals	84.197	67						
Total	94.986	68						
Parameter Estimates (Coefficients)								
95% Confidence Intervals								
Names	Estimate	SE	Lower	Upper	B	Df	T	p
(Intercept)	4.986	0.135	4.716	5.255	0	67	36.942	<0.001
dimension 5, 19	1.243	0.424	0.396	2.089	0.337	67	2.93	0.005

In this study, although the R^2 values may appear modest, they are considered acceptable within the scope of educational and applied research. As Kline [32] points out, in applied settings, methodological decisions are often guided by theoretical coherence and practical relevance rather than strict statistical conventions. Moreover, Cohen [31] provides benchmarks for interpreting R^2 values, suggesting that 0.02, 0.15, and 0.35 represent small, medium, and large effect sizes, respectively. Accordingly, the observed R^2 values (e.g., 0.209 and 0.11) fall within a medium range and provide meaningful insights into the relationship between prompt strategies and response quality. In addition, the use of a six-point scale in this study supports treating the variable as continuous, as research suggests that

Likert-type scales with five or more points can be analyzed as continuous variables without compromising validity [33].

4.3. Description of Simple and Complex Strategies

To address Objective 3, a Rasch analysis [34,35] with a dichotomous model was conducted in Jamovi. The analysis was performed separately for each dimension.

To achieve Objective 3, which aimed to characterize both simple and complex prompting strategies, the responses were first coded dichotomously based on the presence or absence of correct strategy use. Subsequently, the Rasch model was selected and applied for the analysis, as it allows for the estimation of item difficulty by transforming categorical (dichotomous) responses into an interval-level logit scale. This enables an objective comparison of item difficulty across the set.

Contextual Strategies. The MADaQ3 (mean of absolute values of centered Q3) was 0.124, and the p-value exceeded the threshold for significance ($p = 0.375$), thereby satisfying the assumption of local independence [36]. Additionally, all Q3 correlations were below 0.3. CT1 and CT2 were categorized as moderately difficult, with 59% and 47% of participants, respectively, providing correct responses. RBT21 was considered difficult, as only 31% of participants responded correctly (see Table 8).

Table 8. Item statistics—dichotomous model for contextual strategies and Q3 correlation matrix.

	Item Statistics					Q3 Correlation Matrix			
	Proportion	**Measure**	**S.E. Measure**	**Infit**	**Outfit**		**CT2**	**CT1**	**RBT21**
CT2	0.478	0.184	0.355	0.771	0.624	CT2	—		
CT1	0.594	−0.83	0.358	0.81	0.582	CT1	0.004	—	
RBT21	0.319	1.615	0.37	0.905	0.809	RBT21	−0.277	−0.274	—

Note. Infit= Information-weighted mean square statistic; Outfit = Outlier-sensitive means square statistic.

Purpose-Related Strategies. The MADaQ3 value was 0.00, with a corresponding p-value greater than 0.05 ($p = 1.000$), thereby satisfying the assumption of local independence. All Q3 correlations remained below the 0.3 threshold. Techniques CT5 and RBT19 were classified as difficult, with only 21% and 14% of participants, respectively, providing correct responses (see Table 9).

Table 9. Item statistics—dichotomous model for purpose-related strategies and Q3 correlation matrix.

	Item Statistics					Q3 Correlation Matrix		
	Proportion	**Measure**	**S.E. Measure**	**Infit**	**Outfit**		**CT5**	**RBT19**
CT5	0.217	2.05	0.363	0.968	0.963	CT5	—	
RBT19	0.145	2.8	0.414	0.975	0.954	RBT19	−0.263	—

Note. Infit = Information-weighted mean square statistic; Outfit = Outlier-sensitive means square statistic.

5. Discussion

This study aimed to explore the relationship between prompting techniques written in Spanish and the responses generated by a large language model (LLMs), specifically ChatGPT (free version 3.5), as employed by higher education students. Additionally, the study examined how these techniques clustered into emergent strategies and assessed their overall impact on ChatGPT's output quality.

To achieve this purpose, the present research aimed to fulfill the following three key objectives: (i) to identify the emerging prompt strategies employed by higher education students; (ii) to examine the relationship between the types of techniques used

to formulate prompts and the quality of the responses generated by the large language model (LLM); and (iii) to describe both simple and complex strategies observed in students' prompt-writing processes.

From the literature review, we preliminarily proposed two categories of techniques: circumstantial and request-based. Circumstantial techniques were defined as those that provide context and structure to the prompt, forming a foundational layer that guides the interaction. Conversely, request-based techniques refer to those explicitly designed to articulate the query or task directed to the LLMs. For this study, 10 circumstantial techniques and 11 request-based techniques were identified and defined (see Tables 1 and 2). This categorization served as a practical synthesis of existing research, offering a structured framework that was applied to analyze the students' prompt-writing practice.

The key contribution of this study lies in the evidence it provides regarding the impact of prompt engineering strategies on the quality of responses generated by ChatGPT in Spanish. Two emergent strategies were identified, each clustering distinct prompt engineering techniques systematically identified in this research: (i) Guide Contextualization Strategy (GCS) and (ii) Specific Purpose Strategy (SPS). The GCS clusters techniques that focus on providing an informative framework to structure the language model's response generation, ensuring that the AI operates within a well-defined context. In contrast, the SPS group's techniques aimed at directing the model's output by explicitly stating the underlying purpose of the request, thereby shaping the AI's response toward a specific goal. The characteristics and distinctions of each strategy are detailed in the following sections.

Three techniques were clustered into the group of contextual strategies (GCS). The linear regression analysis revealed that CT1 (Context), CT2 (Ambiguity Reduction), and RBT21 (Request with Context) had a significant effect on the quality of responses generated by the model, $F(1, 67) = 17.7$, $p < 0.001$, $R^2 = 0.209$, $\beta = 0.457$. These findings indicate that incorporating contextual information into prompts enhances the accuracy and relevance of the generated outputs.

Although the R^2 value of 0.209 may appear modest, it aligns with the medium effect size threshold proposed by Cohen [31], who defined values of 0.02, 0.15, and 0.35 as small, medium, and large, respectively. In the context of educational and behavioral research, such effect sizes are considered meaningful, particularly when dealing with complex phenomena influenced by multiple factors. Therefore, the conclusion that incorporating contextual information into prompts enhances the accuracy and relevance of the generated outputs remains valid, as it is supported by a statistically significant and practically relevant association.

In addition, the analysis of tetrachoric correlations revealed a strong association between CT1 (Context) and CT2 (Ambiguity Reduction), rt = 0.933, suggesting that these techniques are frequently employed together. CT1 also demonstrated positive correlations with CT5 (Goal or Purpose Indicator), rt = 0.680, and RBT21 (Request with Context), rt = 0.801, indicating that students who apply contextual strategies tend to incorporate elements that clarify the purpose of their prompts.

In contrast, RBT14 (Open-Ended Question Request) showed strong negative correlations with CT1, rt = −0.842, and CT2, rt = −0.866, suggesting that this technique may reduce the likelihood of using clearly contextualized prompting strategies. A possible interpretation is that, in some cases, when students formulate entirely open-ended questions without anchoring them to specific topics or goals, they may overlook key contextual elements needed to guide the generative model effectively. This could result in prompts that lack precision or relevance, potentially affecting the quality and specificity of the AI-generated responses. These findings underscore the importance of fostering a balance between openness and contextual framing in the development of prompt engineering skills.

These findings are consistent with previous research that has identified the use of contextual information as a crucial factor in generating more accurate responses from LLMs. For example, Eager and Brunton [16] highlight the importance of providing a clear and structured context to optimize interactions with LLMs. In this sense, the clarity and specificity of prompts are presented as key factors in obtaining more accurate outputs. Nonetheless, while previous studies have examined these strategies in expert-designed environments, our research focused on the spontaneous interactions between students and ChatGPT, offering a closer perspective on the real-world use of these tools in higher education contexts.

On the other hand, the SPS (Strategies with Purpose Specification) cluster grouped two techniques: CT5 (Goal or Purpose Indicator) and RBT19 (Request with Purpose), both of which demonstrated a significant association with response quality, $F(1, 67) = 8.58$, $p = 0.005$, $R^2 = 0.114$, $\beta = 0.337$. However, the observed effect was smaller than that of the GCS cluster, suggesting that while explicitly articulating the intended purpose in prompts can improve interactions with generative AI, its effective implementation may require a higher level of linguistic proficiency and a deeper understanding of LLMs' operational principles.

These results are consistent with the systematic review conducted by Lee and Palmer [37], who emphasize the importance of users clearly defining their objectives when interacting with AI to maximize effectiveness. Nonetheless, their study revealed that students encounter difficulties in implementing these strategies, suggesting the need for more specialized training in prompt engineering.

Building on this discussion, Figure 3 presents a synthesis of the findings from this study, summarizing the relationships between the identified strategies and their impact on response quality.

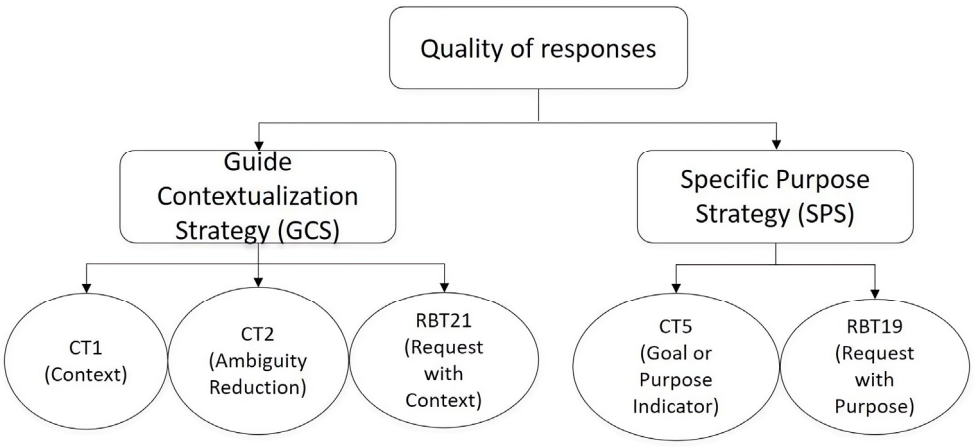

Figure 3. Classification of strategies associated with response quality.

The findings of this study indicate that the techniques grouped within GCS and SPE align with certain prompt engineering strategies documented in the literature. In particular, the techniques classified under GCS share multiple characteristics with the few-shot prompting strategy [21,38,39]. According to Walter [40], incorporating examples and contextual information within a prompt can significantly enhance the quality of responses generated by LLMs. Additionally, Henrickson and Meroño-Peñuela [41] highlight that prompts enriched with prior contextual information are more likely to generate meaningful responses. This observation aligns with the findings of this study, particularly regarding the impact of CT1 (context) and CT2 (ambiguity reduction) on response accuracy. These results suggest that providing a clear contextual framework is a crucial factor in interactions with AI models—an insight that has been consistently identified in few-shot learning research.

Conversely, the techniques grouped under SPS align with strategies that emphasize the explicit formulation of purpose within the prompt. Previous studies have indicated that clearly defining the user's objectives can enhance interactions with LLMs [16]. In particular, Lee and Palmer [37] stress the importance of users establishing well-defined objectives when interacting with AI models. However, their findings also suggest that students struggle to implement these strategies effectively without prior AI literacy in prompt engineering. This observation is consistent with the results of the present study, where SPE demonstrated a weaker effect on response quality compared to GCS, reinforcing the notion that explicit purpose formulation alone may be insufficient without a deeper understanding of how LLMs function.

Furthermore, the literature has identified persona prompting as an effective technique for enhancing interactions with LLMs [11,42]. Although this technique was not explicitly analyzed in the present study, its potential relationship with RBT21 (request with context) warrants further investigation. This is because RBT21, like persona prompting, incorporates a contextual framework in the formulation of requests, making it similar to the role personalization characteristic of persona prompting.

In summary, the findings of this study provide evidence on the importance of contextualization and clear objective definition in the formulation of prompts in Spanish. These results align with previous research on few-shot prompting, persona prompting, and strategies for structuring objectives in interactions with GenAI. However, while many of these studies have been conducted in controlled environments with expert AI users, this research focused on the spontaneous interactions of higher education students with ChatGPT, offering a more authentic perspective on its real-world use in academic settings.

While this study focuses on the functional aspects of prompt construction—specifically the use of circumstantial and request-based techniques—it is also important to recognize that linguistic features, such as sentence complexity and grammatical mood, significantly influence LLM behavior. These language-level dimensions were addressed in a recent study conducted by our team, which examined how variations in grammatical structure affect the perceived quality of AI-generated responses in Spanish-language prompts [27].

Implications of This Work in the Field of Higher Education

The results of this study have significant implications for the development of AI literacy programs in education. As Holmes [43] suggests, it is essential to recognize that the relationship between AI and education is more complex than might initially appear, with misunderstandings often stemming from a lack of comprehensive research [44].

AI literacy encompasses not only technical knowledge but also the ability to engage effectively with the ethical and societal implications of AI technology. In contemporary classrooms, AI literacy must complement traditional learning approaches, equipping students with essential skills to critically assess, interact with, and leverage AI across various aspects of their lives. As an emerging field with immense potential, AI literacy also faces the challenge of early adoption. The difficulty lies not only in imparting technical proficiency but also in fostering a comprehensive understanding of AI's broader impact—whether social, psychological, or economic [40].

Figure 4 presents a methodological framework for structuring interactions with large LLMs through prompts. This framework illustrates the iterative nature of the process, where students formulate prompts using their grammatical skills and prompt engineering strategies, receive responses from the LLM, and evaluate the quality of those responses. If the output is satisfactory, it is accepted; if not, the process restarts with a reformulation of the prompt.

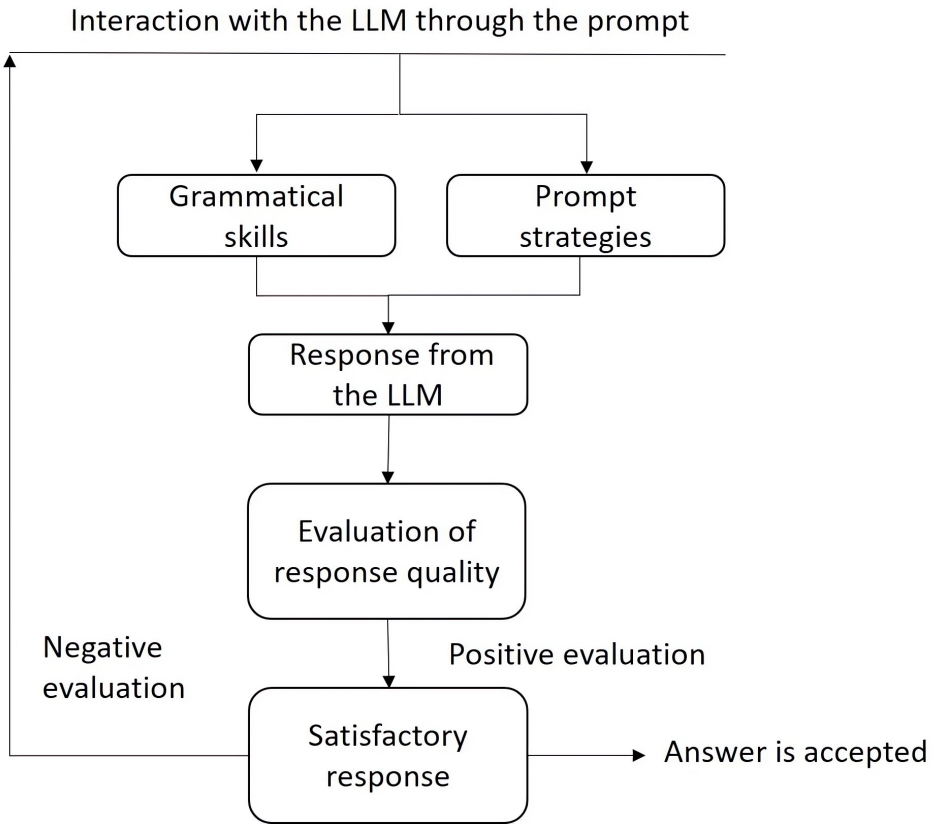

Figure 4. Proposed methodological framework for interacting with LLMs. This figure highlights the role of grammatical skills and prompt strategies in shaping the initial input, followed by AI-generated responses that are evaluated for quality. The horizontal arrows indicate that this process is iterative: users can revise their prompts based on unsatisfactory outputs. The vertical flow underscores the foundational influence of grammatical skills across all stages.

The diagram highlights two foundational components that shape the initial interaction: grammatical skills and prompt strategies. Grammatical performance—particularly in verb selection and syntactic structure—plays a critical role in the clarity and precision of the prompt.

For instance, the use of the subjunctive mood in Spanish, which introduces complexity absent in English, can affect how the model interprets and generates responses [27]. Similarly, compound or subordinate sentence structures can be associated with both the semantic richness and processing of user inputs. As Bozkurt and Sharma [45] note, subtle linguistic variations can significantly impact the relevance of AI-generated content.

Prompt strategies refer to the techniques employed to guide the model toward the intended outcome. These include contextualization, goal setting, and structured formulation. A prompt serves as an input to a generative AI model, guiding the nature and quality of its output [9,10]. As described by Schulhoff et al. [3], well-crafted prompts can improve the coherence, depth, and factual accuracy of LLM outputs. However, as recent studies show, these strategies are still not widely understood by students and require explicit training within AI literacy programs [46–48].

The evaluation phase is essential for determining whether the output meets the user's expectations. This involves critical thinking, the ability to contrast sources, and ethical awareness regarding intellectual property, bias, and reliability [40]. The cyclical nature of the model recognizes that achieving a satisfactory response often requires multiple iterations and refined prompt constructions, reinforcing the importance of deliberate interaction design.

This iterative process also reflects how users in higher education evaluate AI-generated outputs based on a web of interconnected factors, including trust, privacy, and security. While some studies identify trust as a decisive factor in adoption [49], others report it as less influential [6], suggesting that privacy and security concerns may mediate or moderate its effect. In this educational context, a systems-thinking perspective is essential to grasp how students' decisions to adopt and engage with GenAI tools are shaped by multiple, interrelated influences [50].

The findings of this study could have meaningful pedagogical implications, particularly for informing the design of educational interventions aimed at supporting novice users in their interactions with GenAI. The proposed framework might serve as a basis for training programs that foster the development of key competencies—such as grammatical precision, effective use of prompting strategies, and critical evaluation of responses. Educators could design scaffolded learning experiences that guide students in the iterative refinement of prompts, encourage experimentation with different prompt types, and promote reflective assessment of AI outputs. Incorporating these elements into AI literacy curricula might contribute to bridging the gap between basic usage and meaningful, responsible engagement with generative AI tools.

It is important to note that, as an exploratory study, the findings should be interpreted as preliminary and hypothesis-generating. Further empirical research, including longitudinal and cross-cultural studies, is necessary to validate and extend the proposed framework.

6. Conclusions

This exploratory study provides evidence on the impact of prompt engineering strategies on the quality of responses generated by ChatGPT (free version 3.5) in Spanish. Analyzing the techniques used by 102 higher education students, two emergent strategies were identified: the Guide Contextualization Strategy (GCS) and the Specific Purpose Strategy (SPS). The first strategy emphasizes the use of contextual information to guide response generation, while the second focuses on the explicit formulation of the request's purpose.

The results indicate that GCS is more strongly associated with response quality, reinforcing the importance of providing a clear contextual framework to enhance the precision and relevance of interactions with GenAI models. On the other hand, the lower impact of SPS suggests that simply stating the purpose of a request is insufficient unless students also possess a deeper understanding of how LLMs function. This highlights a potential gap in AI literacy, where users may struggle to translate well-defined objectives into effective prompt formulations.

These findings align with prior research on few-shot prompting, persona prompting, and structured objective-setting strategies in interactions with GenAI. However, a key distinction of this study lies in its focus on spontaneous student interactions with ChatGPT, as opposed to controlled environments led by AI experts. This methodological approach offers a closer approximation to the real world.

AI usage in higher education, capturing how students naturally engage with LLMs in academic settings. The results suggest that, while context-driven strategies align with established prompting techniques, students may require additional support to refine their use of structured request-based approaches.

From an educational perspective, these findings reinforce the importance of developing AI digital literacy programs that extend beyond technical knowledge of GenAI models and equip students with effective prompt engineering strategies. The methodological framework proposed in this study, encompassing four key dimensions—grammatical skills; prompt strategies; response from the LLM; and evaluation of response quality—serves as a

foundation for future initiatives aimed at fostering critical and effective engagement with GenAI in Spanish-speaking contexts.

As an exploratory study, this research underscores the need for further investigation into the relationship between GenAI and prompt engineering performance, considering factors such as language proficiency, users' disciplinary backgrounds, and the impact of AI digital literacy programs. Additionally, future research should examine how these strategies manifest across different languages and educational settings and assess the effectiveness of interventions designed to enhance students' AI interaction skills. Given the preliminary nature of our findings, we emphasize the importance of replicating this study in broader and more diverse educational contexts to validate the results and explore their applicability at scale. Ultimately, these insights contribute to a growing understanding of how learners engage with AI in academic contexts, reinforcing the need for tailored educational strategies that bridge the gap between technological advancement and pedagogical practice.

7. Limitations

This exploratory study was conducted with a limited sample of higher education students from the Biobío region in Chile. While the findings underscore the relevance of two emerging strategies for interacting with LLMs, further research is necessary to validate these results and evaluate their applicability across diverse educational settings.

Although the proposed framework provides insights into how higher education students interact with GenAI tools in Spanish, it is important to acknowledge the contextual limitations of this study. The findings are based on a non-probabilistic sample from two universities in the Biobío region of Chile and cannot be generalized to all Spanish-speaking learners or higher education contexts without further cross-cultural validation. Future research should include comparative studies across diverse cultural and linguistic settings to assess the broader applicability and adaptability of the framework.

Author Contributions: Conceptualization, R.V.-M. and C.E.C.-S.; data curation, C.E.C.-S.; formal analysis, R.V.-M. and C.E.C.-S.; investigation, R.V.-M. and C.E.C.-S.; methodology, R.V.-M., J.L.C.-S., C.C.-S. and C.E.C.-S.; resources, S.S.-M.-Q.; supervision, J.L.C.-S.; visualization, S.S.-M.-Q.; writing—original draft, R.V.-M., J.L.C.-S., C.C.-S., S.S.-M.-Q. and C.E.C.-S.; writing—review and editing, R.V.-M., J.L.C.-S. and S.S.-M.-Q. All authors have read and agreed to the published version of the manuscript.

Funding: The author R.V.-M. acknowledges support from ANID FONDECYT Postdoc through grant number 3230356. The author C.C.-S. acknowledges support from grant ANID Capital humano Beca Doctorado Nacional Foil 21231752 Project ID 16930.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki, and approved by the Ethics Committee of Universidad Católica de la Santísima Concepción (protocol code 09/2024 and date of approval 14 March 2024).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The original contributions presented in the study are included in the article; further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Opara, E.; Mfon-Ette Theresa, A.; Aduke, T.C. ChatGPT for teaching, learning and research: Prospects and challenges. *Glob. Acad. J. Humanit. Soc. Sci.* **2023**, *5*, 33–40.
2. Haugsbaken, H.; Hagelia, M. A new AI literacy for the algorithmic age: Prompt engineering or educational promptization? In Proceedings of the 2024 4th International Conference on Applied Artificial Intelligence (ICAPAI), Halden, Norway, 16 April 2024; pp. 1–8.

3. Schulhoff, S.; Ilie, M.; Balepur, N.; Kahadze, K.; Liu, A.; Si, C.; Li, Y.; Gupta, A.; Han, H.; Schulhoff, S.; et al. The prompt report: A systematic survey of prompt engineering techniques. *arXiv* **2025**, arXiv:2406.06608.
4. Jo, H.; Bang, Y. Analyzing ChatGPT adoption drivers with the TOEK framework. *Sci. Rep.* **2023**, *13*, 22606. [CrossRef] [PubMed]
5. Al-kfairy, M.; Mustafa, D.; Kshetri, N.; Insiew, M.; Alfandi, O. Ethical challenges and solutions of generative AI: An interdisciplinary perspective. *Informatics* **2024**, *11*, 58. [CrossRef]
6. Almogren, A.S.; Al-Rahmi, W.M.; Dahri, N.A. Exploring factors influencing the acceptance of ChatGPT in higher education: A smart education perspective. *Heliyon* **2024**, *10*, e31887. [CrossRef]
7. Heng, W.N. Adoption of AI Technology in Education Among UTAR Students: The Case of ChatGPT. Ph.D. Thesis, Universiti Tunku Abdul Rahman (UTAR), Perak, Malaysia, 2023.
8. Hadi, M.U.; Al-Tashi, Q.; Qureshi, R.; Shah, A.; Muneer, A.; Irfan, M.; Zafar, A.; Shaikh, M.B.; Akhtar, N.; Wu, J.; et al. A survey on large language models: Applications, challenges, limitations, and practical usage. *TechRxiv* **2023**. [CrossRef]
9. Heston, T.F.; Khun, C. Prompt engineering in medical education. *Int. Med. Educ.* **2023**, *2*, 198–205. [CrossRef]
10. Meskó, B. Prompt engineering as an important emerging skill for medical professionals: Tutorial. *J. Med. Internet Res.* **2023**, *25*, e50638. [CrossRef]
11. White, J.; Fu, Q.; Hays, S.; Sandborn, M.; Olea, C.; Gilbert, H.; Elnashar, A.; Spencer-Smith, J.; Schmidt, D. A prompt pattern catalog to enhance prompt engineering with ChatGPT. *arXiv* **2023**, arXiv:2302.11382. [CrossRef]
12. Liu, P.; Yuan, W.; Fu, J.; Jiang, Z.; Hayashi, H.; Neubig, G. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Comput. Surv.* **2023**, *55*, 1–35. [CrossRef]
13. Giray, L. Prompt engineering with ChatGPT: A guide for academic writers. *Ann. Biomed. Eng.* **2023**, *51*, 2629–2633. [CrossRef] [PubMed]
14. Brown, T.B.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language models are few-shot learners. In Proceedings of the 34th International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 6–12 December 2020; Volume 33, pp. 1877–1901.
15. Schick, T.; Schütze, H. It's not just size that matters: Small language models are also few-shot learners. *arXiv* **2021**, arXiv:2109.07830.
16. Eager, B.; Brunton, R. Prompting higher education towards AI-augmented teaching and learning practice. *J. Univ. Teach. Learn. Pract.* **2023**, *20*, 02. [CrossRef]
17. Sclar, M.; Choi, Y.; Tsvetkov, Y.; Suhr, A. Quantifying language models' sensitivity to spurious features in prompt design. *arXiv* **2023**, arXiv:2310.11324.
18. Dong, G.; Zhao, J.; Hui, T.; Guo, D.; Wan, W.; Feng, B.; Qiu, Y.; Gongque, Z.; He, K.; Wang, Z. Revisit input perturbation problems for LLMs: A unified robustness evaluation framework for noisy slot filling task. In Proceedings of the CCF International Conference on Natural Language Processing and Chinese Computing 2023, Foshan, China, 12–15 October 2023; Springer: Cham, Switzerland, 2023; pp. 682–694.
19. Deng, Y.; Zhang, W.; Chen, Z.; Gu, Q. Rephrase and respond: Let large language models ask better questions for themselves. *arXiv* **2024**, arXiv:2311.04205.
20. Mialon, G.; Dessì, R.; Lomeli, M.; Nalmpantis, C.; Pasunuru, R.; Raileanu, R.; Rozière, B.; Schick, T.; Dwivedi-Yu, J.; Celikyilmaz, A.; et al. Augmented language models: A survey. *arXiv* **2023**, arXiv:2302.07842.
21. Dang, H.; Mecke, L.; Lehmann, F.; Goller, S.; Buschek, D. How to prompt? Opportunities and challenges of zero- and few-shot learning for human-AI interaction in creative applications of generative models. *arXiv* **2022**, arXiv:2209.01390.
22. Knoth, N.; Janson, A.; Leimeister, J.M. AI literacy and its implications for prompt engineering strategies. *Comput. Educ. Artif. Intell.* **2024**, *6*, 100225. [CrossRef]
23. Ekin, S. Prompt engineering for ChatGPT: A quick guide to techniques, tips, and best practices. *TechRxiv* **2023**. [CrossRef]
24. Sawalha, G.; Taj, I.; Shoufan, A. Analyzing student prompts and their effect on ChatGPT's performance. *Cogent Educ.* **2024**, *11*, 2397200. [CrossRef]
25. Fagbohun, O.; Harrison, R.; Dereventsov, A. An empirical categorization of prompting techniques for large language models: A practitioner's guide. *J. Artif. Intell. Mach. Learn. Data Sci.* **2023**, *1*, 1–11. [CrossRef] [PubMed]
26. Long, D.; Magerko, B. What is AI literacy? Competencies and design considerations. In Proceedings of the CHI'20: CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–16.
27. Viveros-Muñoz, R.; Carrasco-Sáez, J.; Contreras-Saavedra, C.; San-Martín-Quiroga, S.; Contreras-Saavedra, C.E. Does the grammatical structure of prompts influence the responses of generative artificial intelligence? An exploratory analysis in Spanish. *Appl. Sci.* **2025**, *15*, 3882. [CrossRef]
28. Ng, T.K.; Leung, J.K.L.; Chu, S.K.W.; Qiao, S.M. Conceptualizing AI literacy: An exploratory review. *Comput. Educ. Artif. Intell.* **2021**, *2*, 100041. [CrossRef]
29. Carretero, S.; Vuorikari, R.; Punie, Y. *DigComp 2.2: The Digital Competence Framework for Citizens*; Publications Office of the European Union: Luxembourg, 2022.

30. Etikan, I.; Musa, S.A.; Alkassim, R.S. Comparison of convenience sampling and purposive sampling. *Am. J. Theor. Appl. Stat.* **2015**, *5*, 1–4. [CrossRef]
31. Cohen, J. *Statistical Power Analysis for the Behavioral Sciences*, 2nd ed.; Routledge: New York, NY, USA, 2013.
32. Kline, R.B. *Principles and Practice of Structural Equation Modeling*, 4th ed.; Guilford Press: New York, NY, USA, 2015.
33. Norman, G. Likert scales, levels of measurement and the 'laws' of statistics. *Adv. Health Sci. Educ.* **2010**, *15*, 625–632. [CrossRef]
34. Wright, B.; Stone, M. *Best Test Design*; Measurement and Statistics: Chicago, IL, USA, 1979; Available online: https://research.acer.edu.au/measurement/1 (accessed on 10 April 2025).
35. Bond, T. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences*, 3rd ed.; Routledge: London, UK, 2015. [CrossRef]
36. Yen, W.M. Effects of local item dependence on the fit and equating performance of the three-parameter logistic model. *Appl. Psychol. Meas.* **1984**, *8*, 125–145. [CrossRef]
37. Lee, D.; Palmer, E. Prompt engineering in higher education: A systematic review to help inform curricula. *Int. J. Educ. Technol. High. Educ.* **2025**, *22*, 7. [CrossRef]
38. Kojima, T.; Gu, S.S.; Reid, M.; Matsuo, Y.; Iwasawa, Y. Large language models are zero-shot reasoners. *arXiv* **2023**, arXiv:2205.11916.
39. Tam, A. What are Zero-Shot Prompting and Few-Shot Prompting. 2023. Available online: https://machinelearningmastery.com/what-are-zero-shot-prompting-and-few-shot-prompting (accessed on 15 April 2025).
40. Walter, Y. Embracing the future of artificial intelligence in the classroom: The relevance of AI literacy, prompt engineering, and critical thinking in modern education. *Int. J. Educ. Technol. High. Educ.* **2024**, *21*, 15. [CrossRef]
41. Henrickson, L.; Meroño-Peñuela, A. Prompting meaning: A hermeneutic approach to optimising prompt engineering with ChatGPT. *AI Soc.* **2025**, *40*, 903–918. [CrossRef]
42. Fotaris, P.; Mastoras, T.; Lameras, P. Designing educational escape rooms with generative AI: A framework and ChatGPT prompt engineering guide. In Proceedings of the 17th European Conference on Games Based Learning (ECGBL 2023), Enschede, The Netherlands, 5–6 October 2023; Academic Conferences and Publishing Ltd.: Reading, South Oxfordshire, UK, 2023; pp. 1–8.
43. Holmes, W. *The Unintended Consequences of Artificial Intelligence and Education*; Technical Report; Education International: Brussels, Belgium, 2023.
44. Miao, F.; Holmes, W.; Huang, R.; Waynel, H. *Artificial Intelligence and Education. Guidance for Policy-Makers*; UNESCO: Paris, France, 2021.
45. Bozkurt, A.; Sharma, R.C. Generative AI and prompt engineering: The art of whispering to let the genie out of the algorithmic world. *Asian J. Distance Educ.* **2023**, *18*, i–vii.
46. Ali, F.; Ibrahim, B.; Ayoub, S.; Ajmal, A.; Tariq, M. Supporting self-directed learning and self-assessment using TeacherGAIA, a generative AI chatbot application: Learning approaches and prompt engineering. *Learn. Res. Pract.* **2023**, *9*, 135–147. [CrossRef]
47. Aaron, L.; Cho, T.; Shehata, A.; Ba, H. AI literacy. In *Optimizing AI in Higher Education*; State University of New York Press: Albany, NY, USA, 2024; pp. 18–23.
48. Zawacki-Richter, O.; Jung, T.; Tang, R.; Hill, L. New advances in artificial intelligence applications in higher education? *Int. J. Educ. Technol. High. Educ.* **2024**, *21*, 32. [CrossRef]
49. Choudhury, A.; Shamszare, H. Investigating the impact of user trust on the adoption and use of ChatGPT: Survey analysis. *J. Med. Internet Res.* **2023**, *25*, e47184. Available online: https://www.jmir.org/2023/1/e47184 (accessed on 5 April 2025). [CrossRef]
50. Al-kfairy, M. Factors Impacting the Adoption and Acceptance of ChatGPT in Educational Settings: A Narrative Review of Empirical Studies. *Appl. Syst. Innov.* **2024**, *7*, 110. [CrossRef]