

Designing an AI Roleplay System for Contextualized Career Guidance in Mexico

Hamza Matadamas¹ and Mario Caudillo^{1,2}

¹Universidad Amerike, Mexico

²CLeaders, Estado de México, Mexico

edomex18@amerike.edu.mx

mcaudillo@amerike.edu.mx

mcaudillo@CLeaders.com

Abstract: Upper secondary education students in metropolitan areas of Mexico face a persistent gap between their initial career expectations and the real conditions that influence university persistence, including commuting time, institutional costs, and family financial constraints. Traditional career guidance tools, typically based on static psychometric questionnaires, do not incorporate these contextual logistical and socioeconomic factors. This paper presents the design process of a ChatGPT-based roleplay system that simulates professional career pathways for students in the Mexican context of 2026. Rather than reporting the long-term effectiveness of the tool on career decision-making, this paper documents how the system was designed, including the iterative prompt engineering process across three major versions, a four-phase conversational architecture that builds an eight-dimension student profile and guides users through a five-stage life simulation, and a small-scale formative evaluation conducted with upper secondary students to determine whether the system could be used as intended. The contribution of this paper is not a claim of vocational effectiveness, but a reproducible methodological account of how a large language model can be configured through prompt engineering to incorporate contextual, socioeconomic, and culturally situated constraints into a career guidance experience. The formative evaluation provided a descriptive overview of changes observed among the participating students and, more importantly from a design perspective, identified a specific usability issue in the data collection procedure. This issue is discussed in relation to the existing literature on technology adoption in career guidance services.

Keywords: Career Guidance, Generative artificial intelligence, Roleplay, Prompt engineering, Conversational AI, Educational technology

1. Introduction

University dropout during the first years of study remains a persistent challenge in metropolitan areas of Mexico, particularly in the eastern region of the State of Mexico, where an increase of more than 5% has been documented over the last decade (Cisneros-Bravo et al., 2023). This phenomenon cannot be explained solely by academic performance, but rather by a deeper mismatch between students' initial career expectations and the actual conditions they face in professional life, including limited local employment opportunities, family financial constraints, and serious deficiencies in public transportation infrastructure. These issues are often the result of insufficient career guidance during upper secondary education.

Although institutional career guidance services are widely available, current approaches have not kept pace with technological and economic changes. Traditional career guidance relies mainly on static psychometric questionnaires and assessments that evaluate abstract abilities while overlooking the contextual factors that influence students' decisions. These standardized approaches typically ignore critical logistical variables such as daily commuting time between home and campus, the affordability of tuition in relation to the family's financial situation, and the influence of the student's immediate social and cultural environment. As a result, students often make decisions based on an idealized and decontextualized view of university life.

The emergence of generative artificial intelligence and large language models (LLMs) offers a new opportunity to transform career guidance by providing personalized and interactive support at scale. However, the current academic literature reveals an important gap: relatively few studies describe, in practical detail, how these conversational systems can be designed, structured, and constrained to operate effectively within specific socioeconomic contexts. Most published research on LLMs in education focuses on measuring learning outcomes rather than documenting the design rationale, prompt architecture, and iterative refinement process that enables a system to behave as intended (Meyer et al., 2023).

To address this gap, this paper presents the design process of an AI-powered roleplay simulation system for contextualized career guidance, developed through prompt engineering using ChatGPT. The objective of this paper is to document, in a transparent and reproducible manner, the system's design methodology, conversational architecture, iterative prompt development process, and the findings of an initial formative

evaluation conducted with upper secondary education students. Rather than demonstrating that the tool improves long-term career decision-making, this paper aims to provide researchers and practitioners with a reproducible account of how to design a culturally situated AI tutor for career guidance.

2. Literature Review

2.1 Career Guidance

Contemporary literature describes career guidance as a continuous process of learning and career management that aims to develop early decision-making skills, contributing both to personal fulfillment and socioeconomic development (Watts & Sultana, 2004). Comparative policy reviews describe career guidance as a public good that requires coordinated institutional support rather than isolated psychometric assessment (Career Guidance and Public Policy, 2004).

Savickas (2015) identifies three paradigms of career intervention: guiding, developing, and designing. The guiding paradigm, which has historically been the dominant approach, matches individual characteristics with occupational categories through standardized assessment tools. The developing paradigm views career choice as a maturation process that unfolds through developmental stages. The more recent designing paradigm considers career construction as a narrative and adaptive process in which individuals build identity and meaning by constructing their own life stories, making it particularly suitable for uncertain and rapidly changing work environments. The traditional career guidance services discussed in Section 1 operate mainly within the guiding paradigm by relying on static assessment tools that separate students' profiles from their immediate socioeconomic and logistical context.

Beyond the counseling process itself, career guidance services increasingly depend on information and communication technologies (ICT) for their delivery. Kettunen and Sampson (2019) conducted a phenomenographic study with career development experts and identified four categories of barriers to the implementation of ICT in career guidance services: inadequate access to technology, inadequate access to information, insufficient user skills and competencies, and inadequate integration of ICT into existing career guidance workflows. This framework is directly relevant to any AI-mediated career guidance tool, as its successful implementation depends not only on the quality of its content but also on the digital competencies of the end users who interact with the platform.

2.2 Game-based Learning and Roleplay

Situated and experiential learning theories suggest that decision-making skills are strengthened through direct participation in realistic scenarios rather than through passive reflection on the results of abstract assessments. Within this perspective, roleplay simulation is recognized as a high-impact educational strategy that allows learners to experience the dynamics of a specific domain within a controlled and risk-free environment. A roleplay-based simulation is typically organized into three stages: an introductory phase that establishes the rules and context, an interactive gameplay phase, and a debriefing phase in which participants critically reflect on the consequences of their decisions.

Serious games and roleplay experiences have been applied in a wide variety of educational and professional training settings, ranging from language learning to civic education. These experiences generally combine a structured decision-making environment with narrated consequences to transform abstract concepts into meaningful learning experiences. This body of research supports the design decision discussed in Section 3, where career guidance is structured as a multi-stage roleplay experience rather than as a one-time questionnaire.

2.3 Generative AI in Education

Large language models (LLMs) introduce specific challenges for educational applications because of the probabilistic nature of their text generation, which requires domain expertise to critically evaluate their outputs (Meyer et al., 2023). Huber et al. (2024) argue that, because interacting with LLMs involves a very low opportunity cost for users, their use in education carries the risk of overdependence, which may unintentionally reduce the active engagement that meaningful learning requires. The authors suggest that, when appropriate instructional scaffolding is provided, prompt engineering can become a playful way of interacting with LLMs, allowing students to actively explore a topic by modifying prompt structures and

constraints instead of simply receiving a single authoritative answer. This perspective directly informed the one-question-per-turn rule adopted in the system described in this paper (Section 3.1).

Empirical implementations of ChatGPT in real educational settings remain relatively limited. Kavadella et al. (2024) conducted a mixed-methods study involving 77 undergraduate dental students divided into a ChatGPT-assisted group and a traditional literature research group. Their study combined quantitative knowledge assessments with qualitative observations of students' interactions with the tool. The inclusion of a comparison group and a sample size suitable for inferential statistical analysis provides an example of a rigorous evaluation framework. However, this is intentionally not the focus of the present paper, whose main contribution lies in documenting the system's design process and prompt engineering methodology (see Section 6).

2.4 Designing AI-based Educational Systems

A growing body of literature emphasizes the sociotechnical nature of prompting rather than treating it as a purely technical skill. Ağırdağ (2025) introduces the concept of prompting literacy, arguing—drawing on Bourdieu's theory of linguistic capital—that differences in people's ability to interact effectively with generative AI systems may reproduce existing social inequalities instead of reducing them. This has direct implications for AI-based career guidance tools designed for students from socioeconomically disadvantaged contexts. If the ability to operate the interface itself—for example, exporting or sharing a conversation—depends on digital or linguistic skills that are unevenly distributed among students, the system risks systematically excluding the very population it aims to support. Similarly, multidisciplinary efforts to establish guidelines for the responsible integration of generative AI into education emphasize the importance of explicitly documenting system design decisions, including the rationale behind prompt construction and the safeguards incorporated into AI-mediated interactions. Such documentation is considered an essential requirement for the responsible adoption of generative AI in educational settings (Symeou et al., 2025).

3. Design Methodology

This section describes how the system was developed before any interaction with students took place.

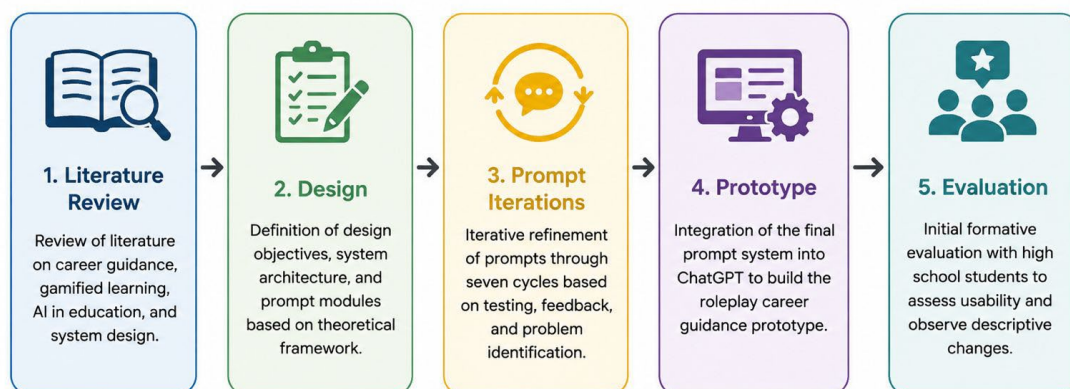


Figure 1: Research process

3.1 Design Objectives

Table 1: Design Objective vs Implementation in the System Prompt

Design objective	Implementation in the System Prompt
Guide student reflection without prescribing a single career choice	Phase 2 provides a Top 5 list of compatible careers based on affinity scores rather than offering a single recommendation.
Make socioeconomic and logistical constraints explicit	Explicit prompt rules require the system to mention real university campuses, commuting time, and cost of living whenever these factors are relevant.

Design objective	Implementation in the System Prompt
Maintain narrative immersion	The simulation is narrated in the second person and uses open-ended questions instead of closed A/B options.
Reduce cognitive overload	The system follows a strict one-question-per-turn rule throughout the interaction.

The one-question-per-turn rule (Objective 4) was not only a usability decision but was also informed by Huber et al. (2024), who argue that low-friction interaction with large language models may encourage overdependence and reduce meaningful engagement with the learning experience. By spacing the interaction into discrete and reflective turns, the system was designed to preserve the student's active role in the decision-making process.

3.2 Prompt Engineering Process

The system prompt was not developed as a single static instruction. Instead, it evolved through a continuous and incremental refinement process consisting of seven iterations. Rather than creating completely different versions of the system, each iteration introduced small cumulative adjustments that addressed specific issues identified while testing the model with ambiguous, inconsistent, or unexpected user inputs.

Two examples illustrate this refinement process. During an early iteration, the model tended to recommend too many career options at the same time, overwhelming students with excessive information. To address this issue, the prompt was modified to explicitly limit the number of career recommendations presented during each interaction. In another iteration, the generated narratives were found to be overly optimistic, portraying professional careers with very little uncertainty or risk and reinforcing the idealized view of university life that this project intended to challenge. As a result, the prompt was revised to strengthen the reality factor and the controlled failure mechanics described in Section 3.1.

The remaining modifications involved smaller refinements, including improvements to formatting, conversational tone, and the structure of the reflection questions. These changes accumulated throughout the design process until reaching the final version used in the formative evaluation, whose functional architecture is described in Section 4. The final version also established the closing instructions that guided students through the conversation export process discussed in Section 5.

Table 2: Prompt modules and purpose

Module	Purpose
System Prompt / Role Instructions	Defines the model as a career guidance counselor and life simulation narrator with expertise in the Mexican labor market of 2026.
Context Prompt (Student Profile)	Collects the eight-dimensional student profile described in Section 4.1.
Narrative Prompt (Scenario/Roleplay)	Generates the list of recommended careers and guides the five-stage life simulation described in Section 4.3.
Reflection Prompt	Presents an open-ended reflection question after each narrated consequence before moving to the next decision.
Closing / Feedback Prompt	Concludes the interaction and provides instructions for exporting the conversation as part of the formative evaluation process.

3.3 Conversational Architecture

The conversational architecture operates as a modular pipeline in which each user input is processed sequentially through the context module (student profile), the persona module (career counselor behavior adapted to the Mexican context of 2026), and the narrative module (simulation rules) before being sent to the ChatGPT engine. The generated response consists of an immersive narrative followed by a single open-ended reflection question that guides the student's next decision.

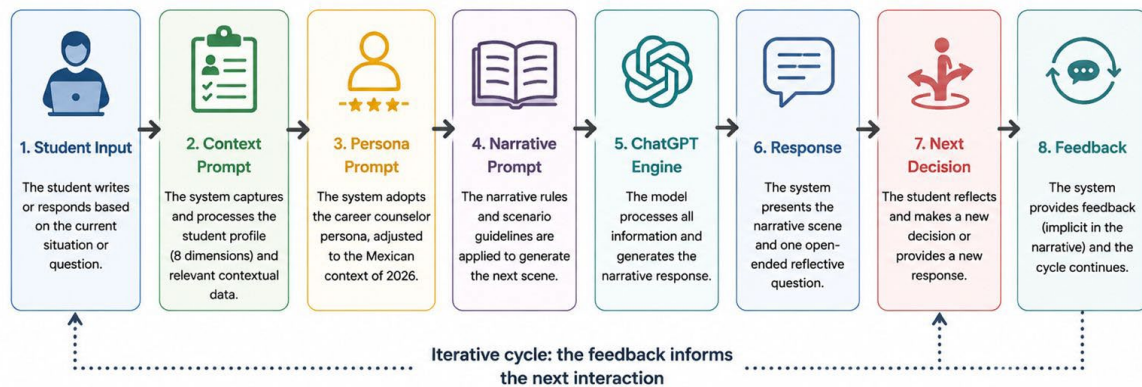


Figure 2: Conversational Architecture

4. System Design

This section describes the functionality of the final system, independently of the design process through which it was developed.

4.1 Student Profile (Phase 1)

The interaction begins with a fixed welcome message, after which the system collects an eight-dimensional student profile, asking one question per conversational turn.

Table 3: Data collected from eight-dimensional student profile

Dimension	Data Collected
1. RIASEC Interests	Self-rating (1–5) for Technology, Helping, Business, Arts, Logic, and Research.
2. Strengths	Up to three strengths selected from Mathematics, Communication, Creativity, Leadership, Technical Skills, and Empathy.
3. Preferences	Up to two preferences selected from Working with People, Data, Creating, Leading, and Research.
4. Aversions	Up to two dislikes selected from Blood, Mathematics, Public Speaking, Computer-Based Work, and Physical Effort.
5. Motivation	Two primary motivations selected from Income, Stability, Freedom, Social Impact, and Creativity.
6. Socioeconomic and Logistical Context	Current location, willingness to relocate, available budget, expected and ideal salary, and preferred work modality.
7. Family Influence	The student's perception of their family's expectations regarding their future career.
8. Career Vision	Open-ended question: "If you could start working in any profession tomorrow, what would you choose?"

4.2 Scenario Generation (Phase 2)

Based on the collected profile, the system generates a RIASEC-inspired compatibility profile together with a Top 5 list of recommended careers. For each recommendation, the system reports a compatibility percentage, a suggested university and campus that match the student's geographical location and financial circumstances (e.g., UNAM FES Acatlán, ITAM Río Hondo, or Tecnológico de Monterrey, Guadalajara Campus), and a feasibility analysis that includes practical logistical information such as estimated commuting time and the level of competition within the profession.

This phase concludes by asking the student which career they would like to explore through the roleplay simulation.

4.3 Roleplay Mechanics (Phase 3)

The selected career is presented as a second-person narrative that unfolds across five cumulative stages, where decisions made in earlier stages influence the events and opportunities encountered later in the simulation.

Table 4: Roleplay stages and mechanics

Stage	Mechanics
1. Admission	The student completes entrance examinations and administrative procedures. Failure to gain admission may require selecting a less preferred alternative career.
2. University	The student faces academic and personal challenges. Prioritizing social life negatively affects academic performance, while combining work and study is portrayed as physically and emotionally exhausting.
3. Professional Internship	The student enters the workplace with limited professional experience. They are portrayed as an inexperienced newcomer who may face demanding schedules and potentially exploitative working conditions.
4. Job Market	The student applies for real job opportunities. Aggressive salary negotiation as an entry-level candidate may result in losing the offer to another applicant, followed by several months of narrated job searching.
5. Adult Professional Life	The simulation explores long-term career outcomes such as promotion, burnout, or entrepreneurship, including the possibility of losing the initial financial investment when starting a business.

After each stage, the system explains the consequences of the student's decisions, describing their impact on finances, physical and mental well-being, and professional network. It then presents one open-ended reflection question before continuing to the next stage (e.g., *"How does this lifestyle make you feel?"* or *"Is this the sacrifice you expected?"*).

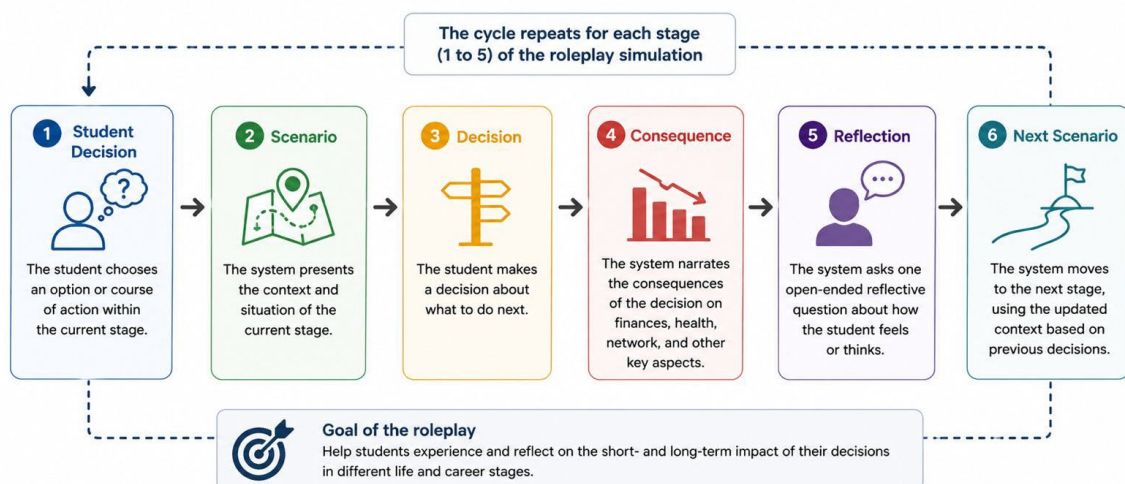


Figure 3: Roleplay Flow

4.4 Reflection and Closing Stage (Phase 4)

Once the simulation reaches a final outcome—successful completion, voluntary withdrawal, or a narrated professional game over—the system thanks the student for participating and asks them to export and share the complete ChatGPT conversation through a linked online form. Step-by-step instructions are provided for generating a compatible sharing link.

As reported in Section 5, this export procedure became the primary usability issue identified during the formative evaluation.

4.5 Formative Usability Evaluation

The formative evaluation was conducted remotely with upper secondary education students from the State of Mexico using a three-stage procedure consisting of an initial questionnaire, a roleplay session, and a post-intervention questionnaire. Given the small, non-random sample available, the results reported in this section are purely descriptive. No inferential statistical analyses were performed, and no claims of effect or association are made based on the percentages derived from a sample of this size.

Table 5: Aspects of formative evaluation.

Aspect	Observation
Initial questionnaire completed	13
Final questionnaire completed (matched responses analysed)	8
Delivery mode	Remote
Primary purpose	Formative evaluation of system usability and conversational coherence, rather than measuring intervention effectiveness

With this methodological limitation explicitly acknowledged, the observed changes between the initial and final questionnaires (n = 8 matched responses) are presented below for descriptive purposes only and should not be interpreted as evidence of an intervention effect.

Table 6: Results

Observed Dimension	Initial Questionnaire	Final Questionnaire
Defined career choice	50%	62.5%
No perceived family pressure	50%	75%
Tolerance for commuting more than one hour	37.5%	12.5%
High understanding (4-5) of institutional costs	37.5%	50%
High perception of institutional location suitability	50%	62.5%
High knowledge (4-5) of the professional field	62.5%	62.5%
High satisfaction (4-5) with current academic situation	50%	62.5%

These values are reported solely as a description of the available sample. With n = 8, each participant represents 12.5 percentage points, meaning that a single response change produces the magnitude of variation shown above. Consequently, no significance tests were performed, and no claim is made that these differences are attributable to the intervention.

The most consistent operational finding was a usability issue associated with the closing procedure described in Section 4.4. Several students submitted incomplete ChatGPT sharing links or links pointing to a single message rather than the complete conversation. Interpreted through the framework proposed by Kettunen and Sampson (2019), this finding corresponds to the category of insufficient user skills and competencies in the use of ICT within career guidance services rather than to a technical failure of the system itself. Viewed through Ağırdağ's (2025) concept of prompting literacy, it also suggests that students with lower levels of digital or linguistic capital may have experienced greater difficulty completing the conversation export process. As a result, the analyzed sample may underrepresent precisely the population that the system was designed to support.

5. Ethics and Consent

This study was conducted as part of the undergraduate course *Research for Video Games* within the Bachelor's Degree in *Artistic Creation for Video Games and Digital Media* at Universidad Amerike, Satélite Campus. No institutional ethics approval was sought because the study was conducted as a classroom research project rather than under a formal institutional research protocol.

Participants were recruited informally through the principal investigator's personal network, including upper secondary education students known to the researcher and, through snowball sampling, their classmates. Participation was voluntary, and students were verbally informed that they were participating in a research experiment. No written informed consent was collected, and no parental or guardian consent was obtained.

The authors recognize this as a significant limitation of the study. The absence of a formal informed consent procedure —particularly given that some participants may have been minors—limits the ethical rigor of the research design. Any future study involving a larger sample should therefore include a formal ethics protocol comprising written informed consent, parental or guardian consent where appropriate, and review by an institutional ethics committee.

6. Discussion

The primary comparison of interest in this study is not between measured outcomes, given the evaluation design described above, but between alternative design decisions. Unlike traditional career guidance approaches based on static psychometric assessments, the proposed system integrates socioeconomic and logistical constraints into a dynamic narrative simulation, positioning career exploration within Savickas' (2015) career design paradigm rather than the traditional guiding paradigm.

Two design lessons emerged from the development process described in Sections 3 and 5. First, limiting the interaction to a single question per conversational turn proved essential for reducing cognitive overload among younger participants while preserving the intended reflective pace. Without this restriction, early prompt versions tended to generate lengthy encyclopedic responses with minimal user engagement, precisely the interaction pattern that Huber et al. (2024) associate with overdependence on large language models.

Second, explicitly incorporating the reality factor and controlled failure mechanics was essential for counteracting idealized expectations about higher education and the labor market. However, the potential emotional impact of these mechanics on adolescent users, discussed in Section 5.1, requires additional ethical consideration before the system can be implemented at a larger scale.

The formative evaluation also revealed an informative limitation. Relying on a built-in feature of a general-purpose consumer platform (the ChatGPT conversation sharing link) as the primary mechanism for data collection introduced a usability barrier that disproportionately affected the least digitally prepared participants. This finding is consistent with previous research on ICT adoption barriers in career guidance services (Kettunen & Sampson, 2019) and inequalities in prompting literacy (Ağırdağ, 2025). Future iterations should therefore replace this mechanism with a purpose-built data collection process requiring a lower level of technical competence.

7. Conclusions

This paper presents the design process, conversational architecture, prompt engineering methodology, and formative evaluation of an AI-powered roleplay system for contextualized career guidance. Rather than claiming improvements in long-term career outcomes, the paper provides a reproducible methodological account of how a general-purpose large language model can be configured through iterative prompt engineering to incorporate economic, geographical, and emotional factors into a career guidance experience.

The formative evaluation suggests that the system is technically usable and capable of producing coherent conversations that consistently follow its intended interaction phases. At the same time, it identifies a specific usability and equity issue in the conversation export procedure that should inform the design of future iterations and larger-scale evaluation studies.

Acknowledgements

We want to acknowledge the technical and financial support of Universidad Amerike Campus Satellite, in the production of this work.

Ethics Declaration: This is to acknowledge that the authors report there are no competing interests to declare.

AI Declaration: The authors used OpenAI's ChatGPT to assist in organizing ideas, for grammar and style editing in english without generating content. Additionally, the Scite platform was employed to perform intelligent searches for relevant, recent papers indexed in Scopus and Web of Science. Outputs from these tools were used solely to inform structural organization and language refinement; all analytic content and interpretations remain the authors' own.

References

- Ağırdag, O. (2025). Beyond prompt engineering: Prompting (I)literacy, linguistic capital, and educational inequality. *Educational Theory*, 76(2), 206–223. HYPERLINK "<https://doi.org/10.1111/edth.70057>" \h<https://doi.org/10.1111/edth.70057>
- Career Guidance and Public Policy. (2004). OECD Publishing. HYPERLINK "<https://doi.org/10.1787/9789264105669-en>" \h<https://doi.org/10.1787/9789264105669-en>
- Cisneros-Bravo, B. E., Rodríguez-Aguilar, R. M., Niño-Membrillo, Y. E., & Cuevas-Rasgado, A. D. (2023). Falta de orientación vocacional como factor en la deserción universitaria. Caso de estudio: zona Oriente del Estado de México. *RIDE. Revista Iberoamericana para la Investigación y el Desarrollo Educativo*. HYPERLINK "<https://mail.ride.org.mx/index.php/RIDE/article/view/1715>" \h<https://mail.ride.org.mx/index.php/RIDE/article/view/1715>
- Huber, S. E., Kiili, K., Nebel, S., Ryan, R. M., Sailer, M., & Ninaus, M. (2024). Leveraging the potential of large language models in education through playful and game-based learning. *Educational Psychology Review*, 36(1), 25. HYPERLINK "<https://doi.org/10.1007/s10648-024-09868-z>" \h<https://doi.org/10.1007/s10648-024-09868-z>
- Kavadella, A., Dias da Silva, M. A., Kaklamanos, E. G., Stamatopoulos, V., & Giannakopoulos, K. (2024). Evaluation of ChatGPT's real-life implementation in undergraduate dental education: Mixed methods study. *JMIR Medical Education*, 10, e51344. HYPERLINK "<https://doi.org/10.2196/51344>" \h<https://doi.org/10.2196/51344>
- Kettunen, J., & Sampson, J. P. (2019). Challenges in implementing ICT in career services: Perspectives from career development experts. *International Journal for Educational and Vocational Guidance*, 19(1), 1–18. HYPERLINK "<https://doi.org/10.1007/s10775-018-9365-6>" \h<https://doi.org/10.1007/s10775-018-9365-6>
- Kuz, A., & Morales, R. (2023). Ciencia de Datos Educativos y aprendizaje automático: un caso de estudio sobre la deserción estudiantil universitaria en México. *Education in the Knowledge Society (EKS)*, 24, e30080. HYPERLINK "<https://doi.org/10.14201/eks.30080>" \h<https://doi.org/10.14201/eks.30080>
- Meyer, J. G., Urbanowicz, R. J., Martin, P., O'Connor, K., Li, R., Peng, P.-C., Bright, T. J., Tatonetti, N. P., Won, K., Gonzalez-Hernandez, G., & Moore, J. H. (2023). ChatGPT and large language models in academia: Opportunities and challenges. *BioData Mining*, 16(1), 20. HYPERLINK "<https://doi.org/10.1186/s13040-023-00339-9>" \h<https://doi.org/10.1186/s13040-023-00339-9>
- Savickas, M. L. (2015). Career counseling paradigms: Guiding, developing, and designing. En P. J. Hartung, M. L. Savickas, & W. B. Walsh (Eds.), *APA handbook of career intervention: Vol. 1. Foundations* (pp. 129–143). American Psychological Association. HYPERLINK "<https://doi.org/10.1037/14438-008>" \h<https://doi.org/10.1037/14438-008>
- Symeou, L., Louca, L. T., Kavadella, A., Mackay, J., Danidou, Y., & Raffay, V. (2025). Development of evidence-based guidelines for the integration of generative AI in university education through a multidisciplinary, consensus-based approach. *European Journal of Dental Education*, 29(2), 285–303. HYPERLINK "<https://doi.org/10.1111/eje.13069>" \h<https://doi.org/10.1111/eje.13069>
- Watts, A., & Sultana, R. G. (2004). Career guidance policies in 37 countries: Contrasts and common themes. *International Journal for Educational and Vocational Guidance*, 4(2–3), 105–122. HYPERLINK "<https://doi.org/10.1007/s10775-005-1025-y>" \h<https://doi.org/10.1007/s10775-005-1025-y>