

Artificial intelligence in Higher Education: Evaluating a Custom GPT for Teaching Planning in a Course on American History

**Inteligencia artificial en la educación superior: evaluación de un GPT
personalizado para la planificación docente en una asignatura de
Historia de América**

Antonio Carrasco-Rodríguez
University of Alicante
antonio.carrasco@ua.es
[0000-0003-3913-0200](tel:0000-0003-3913-0200)

1. Supplementary Material S1. Technical description of the custom GPT assistant and workflow

1.1. Overview of the assistant

The study employed a single custom GPT assistant created within ChatGPT using the Custom GPTs (GPT Builder) environment. The assistant was developed specifically for the undergraduate course *America: History from Colonization to the Present* (BA in History, University of Alicante) and was intended to support both instructional tasks (course planning, materials preparation, assessment design) and students' learning processes (study support and completion of practical assignments).

The assistant was initially created in October 2024 and iteratively refined by the author between October and December 2024. It was subsequently used throughout the teaching period (January–June 2025) and during the resit examination phase (July 2025). A later update developed after the course (June 2025), aimed at extending the assistant with a content-personalization module based on students' academic strengths, corresponds to a different research line and is not part of the configuration evaluated in this study.

1.2. Core configuration and system instructions

The assistant's behavior was governed by a set of system-level instructions defined within the GPT Builder. These instructions established:

1. Pedagogical role: the assistant functioned as a course-specific support tool for History of America, not as a general-purpose historical chatbot.
2. User differentiation: two user profiles were specified—*instructor* and *students*—with differentiated communicative goals, depth of explanation, and pedagogical support.
3. Priority of course materials: the assistant was instructed to consult and prioritize the uploaded course materials before relying on its general training data.

4. Language behavior: responses were to be provided in the language used by the user (Spanish or English).

5. Inclusive and respectful language: the assistant was required to adopt inclusive, non-discriminatory language and to avoid offensive or exclusionary formulations.

6. Task domains: the assistant's functions were explicitly aligned with the ten uses analyzed in the manuscript (teaching planning, materials creation, student training, and assessment support).

7. Query handling: when faced with incomplete or overly broad requests, the assistant was instructed to ask for clarifications or to propose structured thematic exploration paths.

These instructions were progressively refined during the iterative development phase to better align the assistant's outputs with the pedagogical and historiographical aims of the course.

1.3. Knowledge base

A dedicated knowledge base was incorporated into the assistant. At the time of implementation, this consisted of 20 course-specific files (the maximum allowed by the GPT Builder). All documents were in Spanish and included:

- Updated lecture notes (approximately 260 pages), produced through iterative collaboration between the instructor and the assistant and subsequently validated by the instructor.
- Slide decks exported as PDF files, containing topic outlines and conceptual schemata.
- The official course syllabus and teaching guide.
- Supporting instructional documents designed to optimize the use of the assistant (including a user guide, a topic-review criteria document, and task/function sheets for instructor and students).

The purpose of this knowledge base was to ground the assistant's outputs in validated academic content and to minimize reliance on generic model knowledge.

1.4. Enabled capabilities

The assistant was configured with all available capabilities at the time of deployment, including:

- Web browsing,
- File analysis,
- The canvas editing environment,

- The code interpreter/data analysis tool.

Image generation was not enabled and, depending on the platform version, may not have been available at the time of creation.

1.5. Model configuration and technical limitations

The GPT Builder environment does not provide direct access to low-level generation parameters such as temperature, top-p sampling, or explicit token limits, nor does it allow the export of run-level logs. As a result, these parameters cannot be reported with precision.

Accordingly, this study documents only those configuration elements that are directly controllable within the platform, namely system instructions, the knowledge base, and enabled capabilities. Control over output specificity, rigor, and pedagogical alignment was implemented indirectly through:

- Explicit instructional constraints (e.g., prioritization of course materials),
- Task-specific prompt design,
- Iterative refinement cycles under human supervision.

This limitation affects strict computational reproducibility but reflects the practical constraints of deploying commercial generative AI tools in real educational contexts.

1.6. Human oversight and iterative workflow

The assistant was used as a support tool rather than as an autonomous agent. All outputs intended for teaching or assessment purposes—such as lecture materials, activity guides, rubrics, examination questions, and evaluation instruments—were systematically reviewed, validated, and edited by the instructor prior to use.

The overall workflow followed iterative cycles of prompt formulation, output generation, instructor evaluation, and revision. These cycles served both as a quality assurance mechanism and as a means of progressively tuning the assistant's responses to the pedagogical, epistemological, and ethical goals of the course.

1.7. Relationship with the main manuscript

This supplementary material provides technical and procedural detail that complements the condensed description included in Section 3.X of the main manuscript. The inclusion of this appendix aims to enhance transparency and to facilitate informed interpretation and contextual replication of the study's findings, while maintaining an appropriate balance between methodological rigor and textual economy in the main article.

2. Supplementary Material S2. Representative prompts used in the study

This supplementary material presents a set of representative prompts corresponding to the ten instructional uses of the custom GPT assistant analyzed in the study. The prompts included here are representative examples of the prompts used for each instructional purpose. They illustrate the typical structure and intent of the instructions employed, rather than providing exhaustive interaction logs. This approach balances transparency and replicability with the practical constraints of iterative, instructor-mediated use of generative AI tools in real educational contexts.

All prompts were designed to operate in conjunction with the assistant's system-level instructions and course-specific knowledge base, and all outputs generated from these prompts were subject to human supervision and validation prior to pedagogical use.

Use 1. Review and adjustment of the course syllabus

You are acting as a teaching assistant specialized in Early Modern History. Review the following course syllabus. Assess its internal coherence, alignment between learning objectives, contents, and assessment criteria, and identify possible redundancies or gaps. Propose specific improvements while preserving the historiographical approach and institutional constraints of the course.

Use 2. Design of practical learning activities

Propose three practical learning activities for an undergraduate course in History of America. The activities should be designed respectively around: (a) gamification strategies, (b) role-playing dynamics, and (c) the guided use of the custom GPT assistant. For each activity, specify learning objectives, student tasks, and expected learning outcomes.

Use 3. Structuring of theoretical content (pre-writing phase)

Using exclusively the uploaded course materials, propose a detailed content structure for a section on Inca society. The structure should include thematic blocks, key concepts, and analytical perspectives suitable for undergraduate students. This outline will serve as a preparatory step prior to drafting a preliminary written text.

Use 4. Critical evaluation of historiographical content

Provide a rigorous academic evaluation of the following text on the conquest of the Aztec state. Assess its historical accuracy, conceptual clarity, use of terminology, and interpretative balance. Identify strengths, weaknesses, and possible biases, and suggest concrete improvements.

Use 5. Bibliographic recommendation

Propose a selection of key academic works in English on the exploration of the southern regions of what is now the United States. For each reference, briefly justify its relevance and historiographical contribution. Avoid fabricated or non-verifiable sources.

Use 6. Creation of complementary teaching materials

Design a proposal for complementary teaching materials to support an explanation of Hispanic colonial society. The proposal should include both textual and visual materials,



specify their pedagogical purpose, and indicate how they complement the core lecture content.

Use 7. Autonomous student training

Create a comparative table contrasting the Inca and Aztec empires. The table should include political organization, economic structures, social hierarchy, religion, and territorial control, and be suitable as a study aid for undergraduate students.

Use 8. Design of assessment instruments

Generate an exam set based on the uploaded course materials consisting of: (a) 20 multiple-choice questions, (b) two short-answer questions, (c) one schematic or outline-based question, and (d) one long essay question. Ensure alignment with the course learning objectives and an appropriate range of difficulty levels.

Use 9. Design of analytic grading rubrics for advanced tasks

Design an analytic grading rubric to evaluate a student assignment in which students are required to create an advanced prompt simulating a conversation with a historical character. The rubric should assess historical accuracy, contextualization, creativity, and critical use of generative AI.

Use 10. Evaluation of long essay responses

Create an analytic grading rubric to assess a long essay question requiring historical argumentation. The rubric should include criteria related to factual accuracy, argumentative coherence, use of historical concepts, and critical interpretation.

The representative prompts presented in this supplementary material correspond directly to the ten instructional uses analyzed in the Results section of the main manuscript. Their inclusion aims to enhance methodological transparency and to facilitate informed interpretation of the study's findings, while acknowledging that the actual instructional workflow involved iterative prompt refinement and continuous human oversight.