Technological-Educational Journal DOCENTES 2.0
LA2017000128 | ISSN: 2665-0266 | ISNI: 0000 0005 0409 1664 | Ringgold ID 608948 | ROR: 037jqrk16
RTED
Vol. 19 (2), 2026
June – November 2026
145-155

ARTICLE

Generative Artificial Intelligence in University Instructional Design

Inteligencia Artificial Generativa en el Diseño Instruccional Universitario

Erika Dayana Offerman-Albarrán[1] and Carolina Olave-Rodríguez[2]

[1]Universidad Andrés Bello, Santiago, Chile, email: erika.offerman@uniacc.edu
[2]Universidad Autónoma de Chile, Temuco, Chile, email: di4@uautonoma.cl

Received: March 6, 2026 | **Accepted:** July 6, 2026 | **Published:** November 29, 2026 | **DOI:** 10.37843/rted.v19i2.842

ABSTRACT

The instructional design of online courses faces growing tensions among production demands, pedagogical relevance, and the time available for professional reflection. In this context, the incorporation of generative artificial intelligence models creates opportunities to optimize processes, although it also raises ethical and pedagogical risks that require situated empirical evidence. The aim of this study was to analyze the observed effects of integrating a customized GPT on the efficiency, quality, and pedagogical alignment of instructional design for online courses in higher education. The research was framed within the pragmatic paradigm, using an exploratory sequential mixed-methods design with explanatory integration, a descriptive-explanatory approach, a holistic single-case study design, and a longitudinal time frame. The population consisted of 15 instructional designers and 60 faculty members from Universidad Andrés Bello, selected through purposive criterion sampling. Four instruments validated by expert judgment were applied: a semi-structured interview, a focus group protocol, a satisfaction survey, and a document analysis rubric. Quantitative data were processed in SPSS 28 using descriptive statistics and a paired-samples t-test, while qualitative data were coded in NVivo 14 through reflexive thematic analysis. The results showed a statistically significant 30.4% reduction in the time required to produce instructional deliverables, improvements in five of the six quality criteria, and an overall satisfaction score of 4.3 out of 5. It was concluded that institutional customization of the model may be a relevant factor for strengthening pedagogical relevance, provided that it is articulated with human supervision, institutional criteria, and ethical protocols regarding privacy, algorithmic bias, and authorship.

KEYWORDS: Instructional design, generative artificial intelligence, online courses, higher education, pedagogical alignment.

RESUMEN

El diseño instruccional de cursos en línea enfrentó tensiones crecientes entre la producción, la pertinencia pedagógica y el tiempo disponible para la reflexión profesional. En este contexto, la incorporación de modelos de inteligencia artificial generativa abrió oportunidades para optimizar procesos, aunque también planteó riesgos éticos y pedagógicos que requerían evidencia empírica situada. El objetivo del estudio fue analizar los efectos observados de la integración de un GPT personalizado sobre la eficiencia, la calidad y la alineación pedagógica del diseño instruccional de cursos en línea en educación superior. La investigación se enmarcó en el paradigma pragmático, bajo un método mixto secuencial exploratorio con integración explicativa, un enfoque descriptivo-explicativo, un diseño de estudio de caso único holístico y un corte longitudinal. La población estuvo conformada por 15 diseñadores instruccionales y 60 docentes de la Universidad Andrés Bello, seleccionados mediante muestreo intencional según criterios establecidos. Se aplicaron cuatro instrumentos validados por juicio de expertos: entrevista semiestructurada, guion de grupo focal, encuesta de satisfacción y rúbrica de análisis documental. Los datos cuantitativos se procesaron en SPSS 28 mediante estadística descriptiva y la prueba t de muestras relacionadas; los datos cualitativos se codificaron en NVivo 14 mediante un análisis temático reflexivo. Los resultados evidenciaron una reducción estadísticamente significativa del 30,4 % en el tiempo de elaboración de productos instruccionales, mejoras en cinco de los seis criterios de calidad y una satisfacción global de 4,3 sobre 5. Se concluyó que la personalización institucional del modelo constituyó un factor relevante para fortalecer la pertinencia pedagógica, siempre que se articulara con la supervisión humana, los criterios institucionales y los protocolos éticos sobre privacidad, sesgo algorítmico y autoría.

PALABRAS CLAVE: Diseño instruccional, inteligencia artificial generativa, cursos en línea, educación superior, alineación pedagógica.

How to Cite

Offerman-Albarrán, E. D., & Olave-Rodríguez, C. (2026). Generative Artificial Intelligence in University Instructional Design. *Docentes 2.0 Technological-Educational Journal*, *19*(2), 145-155. https://doi.org/10.37843/rted.v19i2.842

OPEN ACCESS
CC BY NC ND

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

Introduction

The instructional design of online courses faces growing tensions among production demands, pedagogical relevance, and the time available for professional reflection. In this context, the introduction of generative artificial intelligence (GenAI) models creates opportunities to optimize processes, although it also raises ethical and pedagogical risks requiring context-specific empirical evidence. Instructional design has become a key professional practice in online higher education because it connects institutional pedagogical principles with decisions about sequencing, assessment, and technological mediation that shape courses (Branch & Dousay, 2015). In this field, sustained growth in online enrollment increases the demands for quality, relevance, and timely production of learning experiences. At the same time, increasingly diverse student profiles call for more flexible, contextualized designs consistent with institutional educational models. Kumar and Ritzhaupt (2017) show that instructional designers devote a significant proportion of their workday to administrative and coordination tasks, reducing the time available for in-depth pedagogical work.

Instructional design teams face a structural paradox: they are required to produce more courses, with greater personalization and in less time, while having limited tools to meet those demands without compromising academic quality. This tension intensifies when available GenAI tools have been trained on generic corpora that do not always reflect each university's pedagogical frameworks, graduate profiles, or institutional terminology (Zawacki-Richter et al., 2019). Consequently, a gap persists between the stated potential of artificial intelligence (AI) in education and its usefulness in specific institutional settings. This gap may discourage sustained adoption of these technologies and deepen faculty distrust of their pedagogical applications (Luckin et al., 2016).

In response to this problem, the specialized literature proposes complementary approaches to guide the educational integration of artificial intelligence. Holmes et al. (2022) propose developing community-wide ethical frameworks based on transparency, equity, and accountability. From a human-centered design perspective, Alfredo et al. (2024) argue that artificial intelligence tools should amplify designers' professional judgment rather than replace it. FitzGerald et al. (2018) argue that pedagogical personalization requires infrastructure, training, and clear institutional criteria. Philipsen et al. (2019) indicate that the effectiveness of technology in online settings depends on faculty professional development. Nevertheless, empirical evidence on the institutional customization of GPT models, understood as adapting the model using a university's own documentation, remains limited, particularly in Latin American contexts and in designs integrating quantitative and qualitative data.

To address this gap, the general objective of this study was to analyze the observed effects of integrating a customized GPT into the instructional design of online courses at Universidad Andrés Bello (UNAB) during the 2024-1 academic period. The specific objectives were to measure changes in the efficiency of the instructional design process; assess the quality and pedagogical alignment of the deliverables produced with model assistance; and characterize designers' and faculty members' perceptions of the tool's usefulness, limitations, and ethical implications. The study was guided by the following research question: How is the integration of a customized GPT associated with variations in the efficiency, quality, and pedagogical alignment of online course instructional design in higher education?

Methodology

The research was situated within the pragmatic paradigm (Creswell & Plano Clark, 2018), because the research problem combined quantifiable dimensions, such as efficiency, production times, satisfaction levels, and pedagogical alignment scores, with interpretive dimensions related to perceptions, ethical tensions, and the professional meaning attributed to the use of the tool. From this perspective, the study did not seek to privilege a single type of evidence but rather to connect qualitative and quantitative data to develop a context-specific understanding of the observed effects of integrating a customized GPT into online course instructional design.

How to Cite
Offerman-Albarrán, E. D., & Olave-Rodríguez, C. (2026). Generative Artificial Intelligence in University Instructional Design. *Docentes 2.0 Technological-Educational Journal*, *19*(2), 145-155. https://doi.org/10.37843/rted.v19i2.842

OPEN ACCESS
CC BY NC ND

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

An exploratory sequential mixed-methods design with explanatory integration and qualitative predominance was adopted. The first phase made it possible to understand the experiences, assessments, and tensions perceived by instructional designers and faculty members when using the customized GPT. The second phase complemented these findings through descriptive and inferential measurements of production times, satisfaction, and instructional deliverable quality. This sequence was appropriate because participants' perceptions guided interpretation of the quantitative data and helped explain mechanisms associated with efficiency, quality, and pedagogical alignment.

The approach was descriptive-explanatory, as it described observed variations in instructional design processes and deliverables and interpreted factors that may have contributed to these variations in the institutional context under study. The design was a holistic single-case study (Yin, 2018), examining a specific institutional experience as an integrated unit of analysis. The study was longitudinal because it compared information from the 2024-1 academic period, conducted between March and July 2024 with support from the customized GPT, with institutional records from the immediately preceding period, 2023-2, when design processes were undertaken without assistance from the tool.

The population comprised the entire instructional design team at Universidad Andrés Bello involved in implementing the customized GPT during the 2024-1 period, corresponding to 15 designers, and 60 faculty members who participated in tool-assisted online course design. Purposive criterion sampling was applied (Patton, 2015), with a census of instructional designers because all team members directly involved in the experience under analysis were included.

Faculty inclusion criteria were having designed at least one online course with assistance from the customized GPT during the 2024-1 academic period; belonging to different faculties to ensure disciplinary variability; participating in some phase of the instructional design process; and having signed informed consent. Faculty who did not complete the course design process or did not use the tool during the defined period were excluded. Theoretical saturation was reached at interview 12 and verified using the procedure of Guest et al. (2006), with the three subsequent interviews yielding no substantive emerging categories.

The unit of analysis was the institutional instructional design process assisted by the customized GPT. For quantitative comparison, the matched unit consisted of instructional deliverables developed by the same design team in two consecutive academic periods: 2023-2, without model assistance, and 2024-1, with customized GPT assistance. This decision made it possible to compare changes in production times and document quality under similar institutional conditions, without assuming experimental equivalence between the periods.

To ensure methodological traceability of the study's analytical variables, three central variables were operationalized: efficiency, quality, and pedagogical alignment (see Table 1). These variables were defined on the basis of the general objective, the specific objectives, and the criteria of the institutional educational model. Efficiency was understood as variation in the time required and volume of instructional deliverables produced; quality concerned structural consistency and didactic relevance; and pedagogical alignment concerned the consistency of deliverables with the institutional educational model and the use of university-specific terminology.

Table 1

Operationalization of Study Variables.

Variable	Dimension	Indicator	Instrument
Efficiency	Production time	Hours per instructional deliverable	Time records
	Production volume	Deliverables per academic cycle	Institutional records
Quality	Structural consistency	Rubric score (0–5)	Document analysis rubric
	Didactic relevance	Rubric score (0–5)	Document analysis rubric

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

Variable	Dimension	Indicator	Instrument
Pedagogical alignment	Consistency with educational model	Agreement index (0–1)	Expert rubric
	Use of institutional terminology	Frequency of key terms	Document analysis

Note. Institutional educational model of UNAB and the descriptors of Branch and Dousay (2015); authors' own work (2024).

Four complementary instruments were used: a semi-structured interview, a focus group protocol, a satisfaction survey, and a document analysis rubric. The semi-structured interview consisted of 12 open-ended questions organized around four themes: experience of use, perceived efficiency, assessment of quality, and ethical dilemmas. The focus group protocol included seven open-ended questions designed to develop a collective understanding of the possibilities, limitations, and conditions for pedagogical use of the customized GPT. The satisfaction survey consisted of 18 items rated on a five-point Likert scale and grouped into four dimensions: perceived usefulness, ease of use, institutional alignment, and intention to use the tool in the future.

The document analysis rubric was developed by the authors based on the institutional educational model and comprised six criteria scored from 0 to 5: clarity of objectives, internal coherence, disciplinary relevance, use of institutional terminology, alignment with graduate profiles, and didactic feasibility. This rubric was applied to a sample of instructional deliverables produced in the 2023-2 and 2024-1 periods to compare observed changes before and after the introduction of the customized GPT. The documents reviewed included planning matrices, course scripts, and assessment rubrics because they are central documents in the institutional instructional design process.

The procedure comprised four phases. In the first phase, the institutional documentation used to configure the customized GPT was systematically organized, including pedagogical guidelines, curricular guidance, graduate profiles, assessment criteria, and instructional design templates. In the second phase, interviews and focus groups were conducted with designers and faculty members to identify perceptions, opportunities, limitations, and ethical tensions. In the third phase, time records, instructional deliverables, and satisfaction survey responses were collected. In the fourth phase, qualitative and quantitative data were integrated using matrices of convergence, divergence, and interpretive expansion.

The customized GPT was configured with the university's own institutional documentation to align its responses with the educational model, institutional terminology, graduate profiles, and instructional design templates used by academic teams. The tool was used to support draft generation, organize instructional sequences, propose assessment criteria, and review the internal coherence of instructional deliverables. Nevertheless, all model outputs were reviewed, adjusted, and validated by human instructional designers, so the model functioned as a supporting assistant rather than a substitute for professional pedagogical judgment.

Content validity of the instruments was established through expert judgment involving five evaluators: three holders of a doctorate in education and two instructional design specialists. Aiken's V coefficient was applied, yielding an overall value of 0.89, above the threshold of 0.70 recommended by Penfield & Giacobbi (2004). Survey reliability was estimated using Cronbach's alpha, with a value of 0.87, classified as good according to George & Mallery (2003). Reliability of the qualitative analysis was ensured through independent double coding of a 30% subset of the data, yielding a Cohen's kappa of 0.82, classified as substantial agreement (Landis & Koch, 1977).

Quantitative analysis was performed using SPSS version 28. Descriptive statistics, including means, standard deviations, frequencies, and percentages, were used to characterize production times, quality scores, satisfaction levels, and pedagogical alignment indicators. Inferential statistics were also applied using Student's paired-samples t-test to compare production times in the period with the customized GPT, 2024-1, with those of the immediately preceding period without model assistance, 2023-2. Normality of the differences was verified using the Shapiro–Wilk test, with p values > 0.05.

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

The paired-samples t-test was justified because equivalent instructional deliverables produced by the same design team were compared at two consecutive institutional time points. This decision made it possible to estimate temporal variation associated with tool use without attributing strict experimental causality, given the absence of random assignment and an external control group. To strengthen interpretation, quantitative results were analyzed alongside qualitative data from interviews, focus groups, and document review.

Qualitative analysis was performed with NVivo version 14, following Braun & Clarke's (2006, 2019) reflexive thematic analysis method. The procedure involved six phases: familiarization with the data, initial coding, searching for themes, reviewing themes, defining themes, and producing the report. Emerging categories were compared between the researchers and reviewed through an external audit to reduce interpretive biases arising from the authors' professional participation in the case under study.

Data integration was performed through methodological triangulation (Denzin, 2012), connecting quantitative and qualitative evidence, and data triangulation, contrasting information from designers, faculty members, and documentary deliverables. Integration was organized through joint results matrices, equivalent to the joint display proposed by Guetterman et al. (2015), allowing convergences, divergences, and expansions across the different types of evidence to be identified. Three mechanisms were used to control researcher bias: structured reflexivity through a weekly log, external auditing of coding and analyses, and participant validation of preliminary findings.

The study was reviewed and approved by the Ethics Committee of Universidad Andrés Bello. All participants signed informed consent before data collection and were informed of the study's objectives, the voluntary nature of participation, the possibility of withdrawing without consequences, and the measures adopted to protect information confidentiality.

Confidentiality was ensured through alphanumeric participant codes, restricted storage of databases, and removal of personally identifiable information from analytical reports. Because the authors had a professional role in the institutional case under analysis, this dual role was acknowledged as a potential source of bias. Reflexivity procedures, external review, and participant validation were implemented to mitigate this risk. The research was conducted in accordance with the principles of the Declaration of Helsinki and the American Psychological Association (2017) ethical guidelines for research involving human participants.

Results

The main finding was that integration of the customized GPT was associated with a reduction in the time required to produce instructional deliverables and an overall improvement in the pedagogical alignment of the documents reviewed. However, changes were not uniform across all evaluated criteria, since disciplinary relevance decreased slightly during the model-assisted period. Findings were organized according to the study's three specific objectives: changes in instructional design process efficiency, quality and pedagogical alignment of deliverables, and participants' perceptions.

Variation in Design Process Efficiency

Comparison of production times between the period without the customized GPT, 2023-2, and the period with the customized GPT, 2024-1, revealed a statistically significant reduction in the average total time required to produce instructional deliverables. The paired-samples t-test yielded $t(14) = 8.47$, $p < 0.001$, with an effect size of $d = 1.82$. Average total time fell from 25.0 hours to 17.5 hours, representing an overall reduction of 30.4%, as shown in Table 2.

Table 2

Comparison of Production Times by Instructional Deliverable.

Deliverable	Without GPT (h)	With GPT (h)	Reduction (%)
Planning matrix	8.2 (1.4)	5.6 (1.1)	31.7
Course script	12.5 (2.1)	8.9 (1.7)	28.8
Assessment rubrics	4.3 (0.9)	3.0 (0.7)	30.2
Overall average	25.0 (3.2)	17.5 (2.6)	30.4

Note. Values represent the mean (standard deviation) of the hours devoted to each deliverable; authors' own work (2024).

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

The greatest reduction was observed for the planning matrix, which decreased by 31.7%, followed by assessment rubrics at 30.2% and the course script at 28.8%. These results showed a sustained reduction across the three instructional deliverables analyzed. The decrease was more pronounced in documents with a greater structural and organizational burden, whereas the course script continued to require the longest average production time in both periods.

Quality and Pedagogical Alignment of Deliverables

Three independent experts applied the document analysis rubric to a sample of 30 instructional deliverables, distributed across 15 documents from the period without the customized GPT and 15 from the period with the customized GPT. Results showed improvements in five of the six criteria evaluated, with an inter-rater agreement index of 0.78. The largest increase occurred in the use of institutional terminology, followed by internal coherence and alignment with graduate profiles, as shown in Table 3.

Table 3

Average Quality and Pedagogical Alignment Scores for Instructional Deliverables.

Criterion	Without GPT	With GPT	Change
Clarity of objectives	3.8	4.3	+0.5
Internal coherence	3.6	4.4	+0.8
Disciplinary relevance	4.1	3.9	−0.2
Institutional terminology	3.2	4.6	+1.4
Alignment with graduate profiles	3.5	4.2	+0.7
Didactic feasibility	4.0	4.1	+0.1

Note. Scale from 0 to 5. Document analysis using a rubric validated through expert judgment; authors' own work (2024).

The criterion with the largest positive change was institutional terminology, increasing from 3.2 to 4.6 points. Internal coherence rose from 3.6 to 4.4 points, while alignment with graduate profiles increased from 3.5 to 4.2 points. Clarity of objectives also improved, rising from 3.8 to 4.3 points. Didactic feasibility showed a smaller change, from 4.0 to 4.1 points. By contrast, disciplinary relevance decreased from 4.1 to 3.9 points, suggesting a potential limitation of the model in adapting to highly specialized content.

Participants' Perceptions

Thematic analysis of interviews and focus groups identified four analytical categories: cognitive redistribution, disciplinary adaptability, ethical tensions, and training needs. These categories grouped instructional designers' and faculty members' perceptions of using the customized GPT in the online course design process. Coding frequencies showed a greater prevalence of units associated with reducing operational workload, shifting toward reflective tasks, and limitations in technical fields, as shown in Table 4.

Table 4

Emerging Thematic Categories from Qualitative Analysis.

Category	Subcategory	Frequency	Source
Cognitive redistribution	Reduction of operational workload	42	E, FG
	Shift toward reflective tasks	38	E, FG
Disciplinary adaptability	Effectiveness in humanities and health	27	E
	Limitations in technical fields	31	E, FG
Ethical tensions	Data privacy	24	E, FG
	Algorithmic bias	19	FG
	Authorship of deliverables	22	E, FG
Training needs	Critical AI literacy	28	E, FG

Note. E: semi-structured interviews; FG: focus groups. Frequency represents the number of text units coded in NVivo 14; authors' own work (2024).

The cognitive redistribution category contained the largest number of coded units, emphasizing the reduction of operational workload and the shift toward reflective tasks. Disciplinary adaptability yielded differentiated results: participants identified greater usefulness in the humanities and health fields but greater limitations in technical fields. Ethical tensions were expressed as concerns about data privacy, algorithmic bias, and authorship of deliverables generated with model assistance. Finally, the need for training emerged as a recurrent condition for guiding critical and pedagogically appropriate use of the tool.

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

Participant Satisfaction

The participant satisfaction survey assessed four dimensions of customized GPT use: perceived usefulness, ease of use, institutional alignment, and intention of future use. Overall satisfaction was 4.3 out of 5, indicating a favorable assessment of the tool in the instructional design process. The dimension with the highest mean score was perceived usefulness, followed by ease of use and intention of future use. Institutional alignment showed a positive result, although with greater variability among participants.

Table 5

Joint Display of Mixed-Methods Results Integration.

Objective	Quantitative result	Qualitative result	Meta-inference
Efficiency	Reduction of 30.4% ($p < 0.001$)	Cognitive redistribution toward reflective tasks	Convergence
Quality and alignment	Increase in 5 of 6 criteria	Better alignment with institutional terminology	Convergence
Perceptions	Overall satisfaction 4.3/5	Emerging ethical tensions	Expansion

Note. The meta-inference follows the criteria of Fetters et al. (2013); authors' own work (2024).

The satisfaction results reported in Table 5 showed a favorable assessment of the customized GPT as a tool supporting instructional design. Perceived usefulness reached the highest mean score, indicating that participants recognized concrete contributions to the organization and production of instructional deliverables. Although institutional alignment was positive, it registered the lowest mean score among the assessed dimensions, suggesting the need to continue adjusting the tool to the university's pedagogical, curricular, and disciplinary criteria.

Mixed-Methods Integration of Findings

Integration of quantitative and qualitative data made it possible to identify convergences and expansions among the results. The reduction in production time aligned with perceptions of cognitive redistribution toward reflective tasks. Improvement in five of the six quality and alignment criteria was associated with a perception of better alignment with institutional terminology. Overall satisfaction of 4.3 out of 5 expanded understanding of the phenomenon because it coexisted with ethical tensions related to privacy, algorithmic bias, and authorship.

Table 6

Mixed-Methods Integration of Results.

Objective	Quantitative result	Qualitative result	Meta-inference
Efficiency	30.4% reduction in average production time	Cognitive redistribution toward reflective tasks	Convergence
Quality and alignment	Increase in five of six assessed criteria	Better alignment with institutional terminology	Convergence
Perceptions	Overall satisfaction of 4.3 out of 5	Emerging ethical tensions	Expansion

Note. Authors' own work (2024).

The mixed-methods integration shown in Table 6 indicates that the customized GPT was valued as a useful tool for reducing production times and strengthening certain aspects of pedagogical alignment. However, it also revealed important limitations in disciplinary relevance and the ethical management of GenAI use. Overall, results indicated that the tool contributed to the instructional design process when used under professional supervision, with defined institutional criteria and human pedagogical review.

Discussion

The findings answer the research question by indicating that the integration of a customized GPT

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

is associated with relevant variations in the efficiency, quality, and pedagogical alignment of online course instructional design, although these vary across disciplinary areas and entail ethical tensions requiring institutional attention. The 30.4% reduction in production time, improvements in five of the six quality criteria, and identification of four stable thematic categories constitute convergent mixed-methods evidence. These results suggest that institutional customization of the model may enhance pedagogical relevance when combined with human oversight, defined curricular criteria, and systematic professional review processes.

The efficiency finding is consistent with the reports of Kumar and Ritzhaupt (2017), who show that instructional designers devote a considerable proportion of their time to operational tasks that could be assisted by technology. The results extend this evidence by showing that reduced time is directed not only toward increasing production volume but also toward cognitive redistribution to tasks of greater pedagogical value. This shift is consistent with Alfredo et al.'s (2024) argument that human-centered design amplifies professional judgment and with Holmes et al. (2019), who argue that AI should free faculty from routine tasks to strengthen pedagogical interaction. Similarly, García-Heredia and Mujica-Sequera (2025) propose that integrating artificial intelligence into research and educational processes should promote more humane technology capable of connecting scientific rigor with the ethical and relational dimensions of academic work.

Improved alignment with the institutional educational model is consistent with Zawacki-Richter et al.'s (2019) warning that educational AI may reproduce external or decontextualized models. Data from this study suggest that this risk can be mitigated when the model is configured with institutional documents, pedagogical guidelines, graduate profiles, and university-specific terminology. However, results also show that institutional alignment alone does not ensure disciplinary relevance across all fields, as that criterion showed a slight decline. This finding qualifies FitzGerald et al.'s (2018) proposal that pedagogical personalization requires robust infrastructure, training, and institutional conditions, since the case analyzed reveals meaningful advances that still depend on expert review and professional mediation by the design team.

The ethical tensions identified concerning privacy, algorithmic bias, and authorship complement the framework proposed by Holmes et al. (2022), who advocate the context-specific development of community-wide ethical frameworks. These findings reinforce that proposal by showing that ethical concerns are not abstract considerations but concrete conditions of professional legitimacy emerging from practice. This evidence also engages critically with Luckin et al. (2016), whose warning about professional judgment being displaced by automation finds a partial counterexample in this study: the human designer's mediation as a critical reviewer of model outputs functions as a mechanism for preserving pedagogical judgment. Therefore, integration of the tool requires protocols for use, bias review, data protection, and explicit criteria concerning academic authorship.

The disciplinary differences identified, reflected in greater perceived effectiveness in the humanities and health fields alongside greater limitations in technical areas, are consistent with Chiu (2024), who documents that generative AI models perform differently according to knowledge domain and degree of specialization of disciplinary vocabulary. This variability indicates that institutional customization should not be understood as a uniform solution for all fields but as a process adaptable to the epistemological, technical, and communicative requirements of each area. The training needs category is also supported by Scherer et al.'s (2019) meta-analysis, which used structural equation modeling applied to TAM to demonstrate that faculty intention to use educational technologies critically depends on perceived usefulness. In the present study, this usefulness is mediated by prior training, trust in the tool, and users' ability to critically evaluate generated responses.

This study makes distinct contributions on three levels. Theoretically, it contributes to the literature on institutional customization of GenAI by documenting a case in which configuration with an institution's own documentation functions as a mechanism for partially mitigating the risk of pedagogical decontextualization. In practical terms, it offers evidence useful to instructional design

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

teams considering adoption of similar tools, particularly institutions seeking to improve efficiency without weakening professional judgment. Methodologically, the use of a joint results integration matrix following Fetters et al. (2013) contributes to the repertoire of mixed-methods integration strategies in educational research in Latin American contexts, a field still insufficiently explored according to Zhai et al. (2021).

The study's limitations should be interpreted in terms of internal and external validity. Regarding internal validity, the novelty effect (Scherer et al., 2019) may have influenced participants' positive assessments, justifying longer-term longitudinal studies. The authors' professional participation in the case analyzed, although addressed through reflexivity, external auditing, and participant validation, is an acknowledged limitation of the design. As for external validity, the single-case study does not permit statistical generalizations, although it does enable analytical transfer to contexts with similar characteristics (Yin, 2018). Finally, the absence of data on student learning outcomes is a substantive limitation because the scope of conclusions is restricted to the instructional design process rather than its ultimate pedagogical impact.

Conclusions

This study makes an empirical contribution to the field of GenAI-assisted instructional design by providing situated evidence on the measurable and perceived effects of institutionally customizing a GPT model in a Latin American higher education context. Its importance lies in shifting the discussion from general promises about artificial intelligence in education toward the analysis of a concrete experience evaluated through integrated quantitative and qualitative data. Findings support the view that model customization is a relevant factor for strengthening the pedagogical relevance of instructional deliverables, provided it is combined with human supervision, explicit institutional criteria, and professional review. In this sense, the value of artificial intelligence is defined not as a replacement for pedagogical judgment but as support for redistributing operational workload and enabling more academically substantive tasks.

Regarding the specific objectives, the results suggest important implications for the future of instructional design in online higher education. For the first objective, reduced production time may guide workflow reorganization toward more efficient processes without separating them from expert review. For the second objective, observed improvements in institutional alignment may encourage the development of more contextualized educational artificial intelligence models, especially at institutions needing to adapt technological tools to their own pedagogical frameworks. For the third objective, the identified ethical tensions will require institutional protocols on privacy, algorithmic bias, transparency, authorship, and academic responsibility. These projections must be approached cautiously because the findings derive from a single case study and do not permit statistical generalizations.

Future research is recommended to develop longitudinal multi-case studies to evaluate the sustainability of the observed effects and their transferability to other institutional contexts. It will also be necessary to connect AI-assisted instructional design with evidence of student learning, academic performance, educational experience, and the quality of pedagogical interaction. Quasi-experimental or experimental designs are also suggested to estimate more accurately the influence of these tools on educational processes. Finally, future research agendas should integrate pedagogical, ethical, legal, and technological dimensions to consolidate the use of customized GenAI models as a rigorous, context-sensitive, and socially responsible field of study.

Acknowledgments

The authors thank the instructional design team and faculty members of Universidad Andrés Bello who participated in this research, as well as the expert evaluators who contributed to instrument validation.

Declaration of Conflicts of Interest

The authors declare that there are no conflicts of interest related to the conduct of this study or the

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

interpretation of its results. They also state that they have no personal, academic, or financial relationships that could influence the development or findings of the research.

Funding Statement

The authors declare that this research received no funding from public, private, or commercial institutions and was carried out using their own resources, which guarantees independence in the research process.

Ethics Statement

The study was reviewed and approved by an institutional ethics committee. All participants signed informed consent before data collection, and confidentiality was ensured through alphanumeric coding, secure information storage, and removal of personally identifiable data. The research was also conducted in accordance with the principles of academic integrity, voluntary participation, and respect for participants' privacy.

Statement on the Use of Artificial Intelligence

The authors declare that a customized GPT model was used as an instructional design support tool and constituted the subject of analysis in this study. All AI-generated content was reviewed and validated through human oversight, ensuring that academic, methodological, and ethical responsibility remained with the researchers.

References

Alfredo, R., Echeverria, V., Jin, Y., Yan, L., Swiecki, Z., Gašević, D., & Martinez-Maldonado, R. (2024). Human-centred learning analytics and AI in education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 6, Article 100215. https://doi.org/10.1016/j.caeai.2024.100215

American Psychological Association. (2017). Ethical principles of psychologists and code of conduct. https://www.apa.org/ethics/code

Branch, R. M., & Dousay, T. A. (2015). *Survey of instructional design models* (5.ª ed.). Association for Educational Communications and Technology.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, 3(2), 77-101. https://doi.org/10.1191/1478088706qp063oa

Braun, V., & Clarke, V. (2019). Reflecting on reflexive thematic analysis. *Qualitative Research in Sport, Exercise and Health*, 11(4), 589-597. https://doi.org/10.1080/2159676X.2019.1628806

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence*, 6, Article 100197. https://doi.org/10.1016/j.caeai.2023.100197

Creswell, J. W., & Plano Clark, V. L. (2018). *Designing and conducting mixed methods research* (3.ª ed.). SAGE Publications.

Denzin, N. K. (2012). Triangulation 2.0. *Journal of Mixed Methods Research*, 6(2), 80-88. https://doi.org/10.1177/1558689812437186

Fetters, M. D., Curry, L. A., & Creswell, J. W. (2013). Achieving integration in mixed methods designs: Principles and practices. *Health Services Research*, 48(6 Pt 2), 2134–2156. https://doi.org/10.1111/1475-6773.12117

FitzGerald, E., Kucirkova, N., Jones, A., Cross, S., Ferguson, R., Herodotou, C., Hillaire, G., & Scanlon, E. (2018). Dimensions of personalisation in technology-enhanced learning: A framework and implications for design. *British Journal of Educational Technology*, 49(1), 165–181. https://doi.org/10.1111/bjet.12534

García-Heredia, N. B., & Mujica-Sequera, R. M. (2025). Entre códigos y corazones: Investigación e inteligencia artificial hacia una tecnología más humana. *Revista Docentes 2.0*, 18(2), 337-345. https://doi.org/10.37843/rted.v18i2.724

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4.ª ed.). Allyn and Bacon.

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods*, 18(1), 59-82. https://doi.org/10.1177/1525822X05279903

Guetterman, T. C., Fetters, M. D., & Creswell, J. W. (2015). Integrating quantitative and qualitative results in health science mixed methods research through joint displays. *Annals of Family Medicine*, 13(6), 554-561. https://doi.org/10.1370/afm.1865

Holmes, W., Bialik, M., & Fadel, C. (2019). *Artificial intelligence in education: Promises and implications for teaching and learning*. Center for Curriculum Redesign.

Holmes, W., Porayska-Pomsta, K., Holstein, K., Sutherland, E., Baker, T., Buckingham Shum, S., Santos, O. C., Rodrigo, M. T., Cukurova, M., Bittencourt, I. I., & Koedinger, K. R. (2022). Ethics of AI in education: Towards a community-wide framework. International Journal of Artificial Intelligence in Education, 32(3), 504–526. https://doi.org/10.1007/s40593-021-00239-1

RTED
Vol. 19 (2), 2026
June – November 2026
145-155

Kumar, S., & Ritzhaupt, A. D. (2017). What do instructional designers in higher education really do? *International Journal on E-Learning*, 16(4), 371-393. https://go.docentes20.com/ry65

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, 33(1), 159-174. https://doi.org/10.2307/2529310

Luckin, R., Holmes, W., Griffiths, M., & Forcier, L. B. (2016). *Intelligence unleashed: An argument for AI in education*. Pearson Education.

Patton, M. Q. (2015). Qualitative research & evaluation methods: Integrating theory and practice (4th ed.). SAGE Publications.

Penfield, R. D., & Giacobbi, P. R. (2004). Applying a score confidence interval to Aiken's item content-relevance index. *Measurement in Physical Education and Exercise Science*, 8(4), 213-225. https://doi.org/10.1207/s15327841mpee0804_3

Philipsen, B., Tondeur, J., Pareja Roblin, N., Vanslambrouck, S., & Zhu, C. (2019). Improving teacher professional development for online and blended learning: A systematic meta-aggregative review. *Educational Technology Research and Development*, 67(5), 1145-1174. https://doi.org/10.1007/s11423-019-09645-8

Scherer, R., Siddiq, F., & Tondeur, J. (2019). The technology acceptance model (TAM): A meta-analytic structural equation modeling approach to explaining teachers' adoption of digital technology in education. *Computers & Education*, 128, 13-35. https://doi.org/10.1016/j.compedu.2018.09.009

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6.ª ed.). SAGE Publications.

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16(1), Article 39. https://doi.org/10.1186/s41239-019-0171-0

Zhai, X., Chu, X., Chai, C. S., Jong, M. S. Y., Istenic, A., Spector, M., Liu, J. B., Yuan, J., & Li, Y. (2021). A review of Artificial Intelligence (AI) in education from 2010 to 2020. *Complexity*, 2021(1), Article 8812542. https://doi.org/10.1155/2021/8812542

How to Cite

Offerman-Albarrán, E. D., & Olave-Rodríguez, C. (2026). Generative Artificial Intelligence in University Instructional Design. *Docentes 2.0 Technological-Educational Journal*, *19*(2), 145-155. https://doi.org/10.37843/rted.v19i2.842

OPEN ACCESS
CC BY NC ND