This version of the contribution has been accepted for publication, after peer review (when applicable) but is not the Version of Record and does not reflect post-acceptance improvements, or any corrections. The Version of Record is available online at: http://dx.doi.org/10.1007/978-3-032-30015-7_35. Use of this Accepted Version is subject to the publisher's Accepted Manuscript terms of use https://www.springernature.com/gp/open-research/policies/accepted-manuscript-terms

Teachers as Vibe-Coders: Authoring Adaptive Learning Simulators Through Natural Language

Gabrielle Anne Uy, Joseph Benjamin Ilagan, Jose Ramon Ilagan, and Francine De Asis

Business Insights Laboratory for Development, Ateneo de Manila University, Quezon City, Philippines
{gluy,jbilagan,jrilagan}@ateneo.edu
francine.deasis@student.ateneo.edu,
WWW home page: https://www.ateneo.edu/

Abstract. Learning simulations are effective pedagogical tools, but their development often requires technical expertise that limits direct teacher involvement in design. This paper examines how teachers can author adaptive learning simulators through natural-language interaction with AI, positioning educators as "vibe-coders" in teacher–AI co-design processes. Using a design-based research methodology, three teachers iteratively authored and deployed simulators across three undergraduate course contexts: Product Management and Costing, Service Management, and Solutions Architecture. Data were collected from teacher reflections, system interaction traces, and post-session surveys completed by students and teachers. Results indicate that teachers were able to construct functional, classroom-usable simulators through iterative prompt refinement, with domain characteristics shaping authoring effort, simulator behavior, and student sense-making. Students reported positive perceptions of clarity, usability, engagement, and learning support, while teacher reflections identified misalignment and constraint-setting as productive mechanisms for instructional refinement. The paper advances the concept of *adaptive instructional authoring*, re-framing adaptivity in educational systems as a design-time, co-creative process rather than solely a runtime property of learner-facing systems, and discusses design implications for natural-language authoring environments that preserve teacher agency in simulation-based higher education.

Keywords: Learning simulations · Generative AI · Vibe-coding · Natural-language authoring · Adaptive instructional systems · Teacher–AI co-design · Simulation-based learning

1 Introduction

This paper examines how teachers can act as vibe-coders to design adaptive learning simulators without programming, using natural-language prompts and concept models. The aim is to lower barriers to simulation authoring while maintaining fidelity and clarity of core ideas. The work is positioned in HCI at the

intersection of accessible creation tools and simulation-based learning in higher education, informed by evidence on simulations improving understanding when paired with scaffolding and debriefs [5] and by experiential learning cycles that link action, reflection, and abstraction [4]. Building on adaptive instructional paradigms, the study focuses not only on learner adaptivity within simulators but also on teacher adaptivity during authoring, situating vibe-coding within emerging frameworks for adaptive instructional authoring. The platform embodies adaptivity through iterative teacher–AI co-design cycles, in which prompt refinements and system feedback progressively align the simulator's logic with pedagogical intent. This approach extends Adaptive Instructional Systems research [7] by focusing on adaptation during instructional design rather than only during content delivery. The contribution is a practical platform and design considerations that support teacher–AI co-creation for classroom use. This study examines adaptive instructional authoring, where teachers iteratively shape simulator behavior through natural language in collaboration with AI, shifting adaptivity from delivery-time learner modeling to design-time instructional construction.

Unlike prior work on LLM-based authoring tools and Adaptive Instructional Systems, which primarily examine adaptivity during learner interaction or runtime system behavior, this work extends prior framing by emphasizing adaptivity as a property of instructional authoring itself—where teachers iteratively adapt simulator logic and feedback through natural-language interaction with AI during design.

1.1 Objectives and Research Questions

This study is guided by the following questions, emphasizing adaptive instructional authoring and teacher–AI co-design:

RQ1) How can teachers in different domains design and iteratively adapt interactive simulators through natural language without coding?

RQ2) How do domain contexts shape adaptive simulator design, usability, and engagement?

RQ3) How do students and teachers perceive learning, adaptivity, and usability after simulator-based sessions?

2 Review of Related Work

Recent advances in generative artificial intelligence (GenAI) have been driven largely by the emergence of large language models (LLMs), which are neural models trained on massive corpora to generate, interpret, and transform natural language [3]. LLMs exhibit strong capabilities in code generation [9], reasoning, and semantic translation, enabling users to specify complex computational tasks using everyday language rather than formal programming syntax. These models function not only as text generators but as general-purpose cognitive tools that can translate intent into executable artifacts, including software code [9], simulations [6], and interactive systems. As a result, LLMs have begun to lower

technical barriers to software creation and have shifted attention toward human–AI collaboration, prompt design, and the interpretability of model outputs in both professional and educational contexts.

Vibe-coding is a way to instruct LLMs to program or perform development given natural language inputs or *prompts* [11][10]. The concept of vibe-coding has extended into the teaching domain [10][11] and integrating resulting artifacts into a learning management system (LMS) [12].

Design-based research (DBR) [2][8][1] is a suitable research methodology for iteratively bridging the gap between research and practice in formal education.

Prior work on Adaptive Instructional Systems has largely focused on learner-facing adaptivity, where systems adjust content, pacing, or feedback in response to learner models during instruction [7]. In contrast, this study shifts the locus of adaptation to the instructional design process itself. Rather than treating adaptivity solely as a runtime property of educational systems, we examine how teachers iteratively adapt simulator logic, parameters, and feedback structures through natural language interaction with AI during authoring. This design-time adaptivity re-frames adaptation as a co-creative process, where pedagogical intent is progressively negotiated and encoded through teacher–AI interaction, extending adaptive instruction beyond delivery-time personalization to include adaptive instructional construction.

3 Methodology

The study employed a design-based research (DBR) methodology to examine teacher–AI co-design of adaptive learning simulators in authentic classroom contexts. DBR supported iterative refinement of both the system and authoring practices while generating transferable design insights. Three teachers from a private university participated as co-designers, each authoring an interactive simulator for a technology entrepreneurship topic: **(1) Product Management and Costing, (2) Service Management, and (3) Solutions Architecture**. These topics were selected to examine how domain characteristics shape simulator authoring and adaptation.

The vibe-coding platform enabled teachers to create interactive simulators using natural-language prompts rather than programming. It comprised two components: (1) a no-code authoring interface for specifying simulator logic in plain language and (2) an interactive visualization for variable manipulation and inspection of system behavior. Through iterative prompt refinement, teachers externalized domain knowledge into simulators that visualized causal relationships and feedback dynamics.

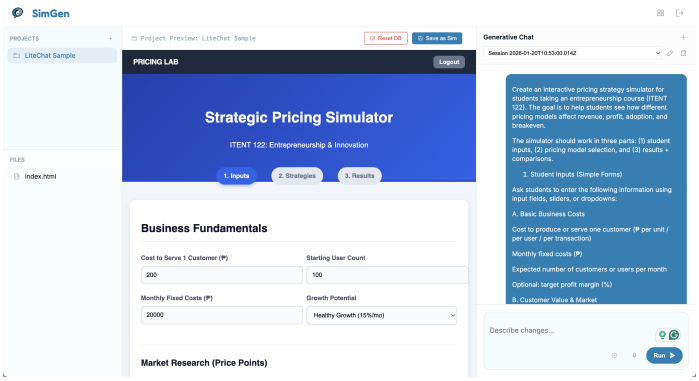


Fig. 1. Simulation generation interface showing teacher-authored natural-language prompts and AI-generated simulator behavior.

Authoring control remained primarily with the teacher throughout the process. Teachers provided free-form natural-language prompts specifying instructional goals, variables, constraints, and feedback logic, while the AI system translated these prompts into executable simulator behavior. Constraints were enforced through iterative prompt refinement rather than hard-coded rules, allowing teachers to narrow scope, clarify assumptions, and adjust simulator feedback based on observed behavior. Prompts were not pre-scaffolded, enabling instructors to author simulators using their own pedagogical language and domain framing.

Each teacher completed iterative authoring cycles involving prompt formulation, inspection of generated simulators, and revision to improve fidelity and interpretability. In this study, adaptivity refers to instructional adaptation achieved at design time through co-authoring between the teacher and AI during these writing cycles. Student feedback was collected only after classroom deployment and used for evaluative purposes; it did not inform further simulator revision within the scope of this study.

One teacher deployed finalized simulators in three live undergraduate classes, where simulators were used as exploratory instructional artifacts through real-time parameter manipulation. Following deployment, students and teachers completed aligned post-session surveys using parallel 5-point Likert-scale items assessing clarity, usability, realism, engagement, and perceived learning support. Teachers additionally provided reflective feedback on authoring effort and instructional fit. No formal assessment of student learning outcomes was conducted.

Data sources included teacher prompt histories and reflection logs, system interaction traces from authoring and classroom use, researcher field notes, and post-session survey responses containing both Likert-scale and open-ended items.

The three research questions were analyzed using complementary quantitative and qualitative approaches. For **RQ1**, teacher reflection responses were

synthesized to examine how instructional intent was articulated through natural-language prompts, where misalignment between intent and system behavior emerged, and how such breakdowns were addressed through iterative prompt refinement and constraint setting. For **RQ2**, student open-ended responses were compared across the three classroom contexts, treating each as a contrastive case to examine how domain characteristics shaped simulator use, perceived clarity, and exploratory behavior. For **RQ3**, descriptive statistics (mean and standard deviation) were computed for Likert-scale survey items, and open-ended responses from both students and teachers were grouped thematically to contextualize these trends. Findings from these analyses are presented narratively in the following section, foregrounding how adaptive instructional authoring and classroom experience co-evolved across domains.

4 Results

This section reports findings aligned with the three research questions. Findings are presented descriptively, emphasizing observed authoring patterns, domain-specific differences, and reported perceptions of usability, engagement, realism, and instructional alignment rather than measured learning outcomes.

4.1 Teacher Authoring Through Natural Language (RQ1)

Across the three instructional domains, teachers reported that natural-language prompting was sufficient to author functional classroom simulators, though authoring typically involved iterative refinement rather than one-time specification. Teachers indicated that early prompts frequently produced simulator behaviors that were overly complex, under-specified, or misaligned with instructional intent, prompting subsequent revisions.

On aligned 5-point Likert-scale items, teachers (n=3) reported strong agreement that they were able to express instructional intent through natural-language prompts (mean = 4.67, SD = 0.58). Teachers also reported that iterative refinement improved simulator alignment over time (mean = 4.33, SD = 0.58). Perceived control over simulator behavior was rated positively but slightly lower (mean = 4.00, SD = 0.00), indicating that some unpredictability persisted even after refinement.

Open-ended teacher reflections described recurring authoring actions, including rewording prompts to clarify assumptions, narrowing the scope of simulator behavior, and adding explicit constraints to reduce unintended outputs. Teachers noted that misalignment between expected and generated behavior was most prominent in early iterations and tended to decrease as prompts became more specific.

4.2 Domain-Specific Patterns in Classroom Use (RQ2)

Student open-ended responses revealed domain-specific patterns in how simulators were used and interpreted across the three classroom contexts. These

contrasts are interpreted as descriptive patterns across contexts rather than statistically tested differences.

In the **Product Management and Costing** context, students (n=10) frequently referenced numerical sensitivity and cause–and–effect relationships among variables such as pricing, costs, and outcomes. Students reported high clarity in understanding how parameter changes affected results (mean = 4.71, SD = 0.46) and indicated that the simulator made relationships between variables easier to observe. Open-ended responses emphasized that small changes in inputs could produce noticeable changes in outcomes, suggesting a focus on quantitative exploration. For example, one student noted that "even small changes in the inputs completely changed the outcome, which made the relationships between variables much clearer."

In the **Service Management** context, student responses (n=15) emphasized tradeoffs and system-level consequences rather than precise numeric relationships. Students described the simulator as supporting reflection on user experience, friction, and service design assumptions. While clarity ratings remained positive (mean = 4.39, SD = 0.67), open-ended responses more frequently referenced interpretive insights, such as reconsidering assumptions about efficiency or security, rather than direct cause–and–effect reasoning.

In the **Solutions Architecture** context, students (n=6) commonly described the simulator as supporting reasoning about automation, reliability, and downstream operational effects. Engagement ratings were high (mean = 4.55, SD = 0.68), but this context also produced more frequent requests for additional explanation to interpret simulator outputs. Students noted that while outcomes were observable, understanding what those outcomes implied required further guidance. As one student described, "I could see the results change, but I needed more explanation to understand what those changes actually meant in practice."

4.3 Student and Teacher Perceptions of Usability and Instructional Alignment (RQ3)

Across all classroom contexts, students (n=31) reported positive perceptions of simulator usability, engagement, and perceived learning support. Students rated ease of interaction highly (mean = 4.65, SD = 0.66) and reported that they understood how to use the simulator during class (mean = 4.61, SD = 0.72). Engagement-related items were also rated positively, including interest in exploring different parameter combinations (mean = 4.61, SD = 0.67).

Perceived learning support ratings indicated that students felt the simulators helped them understand course concepts (mean = 4.55, SD = 0.77) and see relationships between variables and outcomes (mean = 4.52, SD = 0.68). Ratings related to realism were slightly lower but remained positive, with students indicating that the simulators felt realistic (mean = 4.42, SD = 0.81) and reflected real-world situations (mean = 4.42, SD = 0.67).

Teacher responses mirrored these trends. Teachers (n=3) rated simulator usability and classroom integration highly, including ease of classroom use (mean = 5.00, SD = 0.00) and fit within lesson flow (mean = 5.00, SD = 0.00). Teachers

also reported strong perceptions of student engagement (mean = 5.00, SD = 0.00). Similar to student responses, teachers rated realism slightly lower (mean = 4.33, SD = 0.58) than usability and engagement.

Open-ended responses from both students and teachers consistently highlighted the value of experimentation and comparison across scenarios. Students frequently described exploring multiple parameter combinations, while teachers noted that simulator use encouraged discussion and questioning during class sessions.

5 Discussion

This study examined how teachers author adaptive learning simulators through natural-language interaction with AI and how such simulators are perceived and used across different instructional domains. Building on prior work in generative artificial intelligence, vibe-coding, and Adaptive Instructional Systems, the findings suggest that natural-language authoring lowers technical barriers to simulator creation while repositioning adaptivity as a design-time, co-creative instructional process rather than a purely runtime system behavior.

Prior research on large language models has emphasized their capacity to translate natural-language intent into executable artifacts such as software code, simulations, and interactive systems [3][9][6]. The present findings extend this line of work by demonstrating how such intent translation operates in an educational authoring context, where intent is pedagogical rather than purely functional. Teachers in this study used natural-language prompts to express instructional goals, causal relationships, and feedback structures, refining these prompts iteratively as misalignment between intent and system behavior emerged.

Within existing work on vibe-coding, prior studies have shown how natural-language prompting enables instructors and learners to generate instructional artifacts and integrate them into teaching workflows [11][10][12]. The findings here add nuance by showing that when vibe-coding is applied to simulator authoring, the primary challenge is not initial generation but sustained pedagogical alignment. Teachers reported that early simulator outputs were often plausible yet misaligned with instructional intent, requiring prompt refinement, scope narrowing, and constraint specification. This positions vibe-coding not as a replacement for instructional design expertise, but as a medium through which such expertise is iteratively externalized and refined.

The study also contributes to research on Adaptive Instructional Systems by shifting the locus of adaptation from learner-facing runtime personalization to instructional design itself [7]. Whereas prior work has largely conceptualized adaptivity as system-driven adjustment during instruction, the present findings illustrate how adaptation can occur during authoring, as teachers iteratively reshape simulator logic, parameters, and feedback structures through natural-language interaction. This form of design-time adaptivity re-frames adaptation as a co-creative process in which pedagogical intent is progressively negotiated and encoded through teacher–AI interaction.

Domain-specific findings further support this re-framing. In domains characterized by more structured quantitative relationships, teachers reported fewer iterations before achieving acceptable simulator behavior, corresponding with students' strong perceptions of clarity and cause–and-effect understanding. In more interpretive domains, teachers expended greater effort in refining prompts to guide simulator feedback, while students more frequently requested additional explanation to interpret outcomes. This alignment between teacher authoring challenges and student sense-making highlights the importance of interpretability and scaffolding in adaptive instructional authoring.

Across domains, engagement was consistently associated with opportunities for exploration and parameter manipulation, aligning with prior work on experiential learning and simulation-based instruction [5][4]. Importantly, simulators functioned as shared instructional artifacts that supported discussion and reflection rather than replacing instructional facilitation. At the same time, teacher reflections on grounding and unintended system behavior underscore ongoing challenges related to interpretability and accountability in AI-assisted authoring.

These findings suggest that natural-language authoring can support adaptive, simulation-based instruction while maintaining teacher agency, particularly when refinement and constraint-setting remain under instructional control. This perspective re-frames adaptivity as a design-time process embedded in teacher–AI collaboration, extending existing work in adaptive instruction toward a model in which teachers actively construct, rather than merely deploy, adaptive learning artifacts.

6 Conclusion

This paper examined how teachers can author adaptive learning simulators through natural-language interaction with AI, positioning educators as "vibe-coders" in teacher–AI co-design processes. Through a design-based research study spanning three instructional domains, the study demonstrated that teachers can iteratively construct classroom-usable simulators without programming, while actively shaping simulator behavior through feedback and refinement. Rather than treating adaptivity as a system feature applied during instruction, the findings indicate that it can be embedded in the authoring process itself.

6.1 Contributions

This work introduces the notion of adaptive instructional authoring, framing adaptivity in educational systems as a design-time activity shaped through teachers' interaction with AI rather than a property enacted solely during instructional delivery. Through natural-language authoring, teachers iteratively negotiated instructional goals, constraints, and feedback structures, treating simulator construction itself as an adaptive process. In doing so, the study extends

prior work on Adaptive Instructional Systems by shifting attention from learner-facing runtime personalization to the construction of instructional representations during authoring.

At an empirical level, the study shows that natural-language authoring can support the creation of functional, pedagogically aligned simulators across domains with different levels of structure and interpretive complexity. At a conceptual level, it frames vibe-coding not simply as a technique for generating instructional artifacts, but as a co-creative instructional practice through which teacher expertise is articulated, tested, and refined in collaboration with AI.

6.2 Implications

The findings suggest several implications for the design of natural-language authoring systems in education. First, authoring tools should support explicit constraint specification and grounding mechanisms to help teachers manage system assumptions and reduce unintended or misleading behaviors, particularly in interpretive domains. Second, adaptive simulators should incorporate explanation and interpretation layers that help learners translate observable outcomes into meaningful insights, especially when reasoning involves tradeoffs or second-order effects. Third, transparent intermediate representations—such as visible assumptions, causal relationships, or parameter definitions—may support instructional accountability while preserving teacher agency.

For HCI, this work advances research on human–AI co-design by illustrating how adaptive authoring environments can center professional judgment rather than automate it away. The findings emphasize the role of teachers as active co-constructors of instructional models, with AI serving as a responsive and negotiable design partner, rather than casting AI as an autonomous instructional agent.

6.3 Limitations

Several limitations should be considered when interpreting these findings. The study involved a small number of teachers and students from a single institutional context, which limits generalization. Teachers' authoring experiences and perceptions were self-reported, while student interaction data and responses offered evaluative insight into classroom use rather than evidence of system adaptation. Participating teachers also had familiarity with their instructional domains and some exposure to technical concepts, which may have influenced their ease of authoring.

These constraints are consistent with the goals of design-based research, which emphasizes depth of insight and iterative refinement over broad representativeness in early-stage exploration. Accordingly, the findings should be understood as evidence of feasibility and instructional alignment rather than demonstrations of instructional effectiveness across populations.

6.4 Future Work

Future studies could examine learning outcomes through pre–post or comparative designs to better understand how simulator-based instruction authored through natural language influences conceptual understanding. Longitudinal work may also explore how teachers' authoring practices change as simulators are reused, adapted, and shared across courses and cohorts. Additional research could investigate how natural-language authoring platforms support educators with differing levels of technical background, as well as how interface features such as constraint visualization, explanation generation, and authoring feedback shape instructional quality in more complex or interpretive settings.

Acknowledgments. The authors thank the John Gokongwei School of Management, the Ateneo Business Insights Laboratory for Development (BUILD), and the Ateneo University Research Office for their support of this study and its future presentation.

Disclosure of Interests. The authors declare that they have no competing interests that are relevant to the content of this article.

References

1. Anderson, T., Shattuck, J.: Design-Based Research: A Decade of Progress in Education Research? Educational Researcher **41**(1), 16–25 (Jan 2012). https://doi.org/10.3102/0013189X11428813, https://journals.sagepub.com/doi/10.3102/0013189X11428813
2. Brown, A.L.: Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings. Journal of the Learning Sciences **2**(2), 141–178 (Apr 1992). https://doi.org/10.1207/s15327809jls0202_2, http://www.tandfonline.com/doi/abs/10.1207/s15327809jls0202_2
3. Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J.D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A.: Language models are few-shot learners. Advances in neural information processing systems **33**, 1877–1901 (2020), https://proceedings.neurips.cc/paper_files/paper/2020/hash/1457c0d6bfcb4967418bfb8ac142f64a-Abstract.html?utm_source=transaction&utm_medium=email&utm_campaign=linkedin_newsletter
4. Ilagan, J.B., Ilagan, J.R., Rodrigo, M.M.: An agent-based modeling and simulation tool as a learning aid for diffusion of innovations. In: International Conference on Computers in Education (2023), https://library.apsce.net/index.php/ICCE/article/view/4691
5. Kolb, A.Y., Kolb, D.A.: Learning Styles and Learning Spaces: Enhancing Experiential Learning in Higher Education. Academy of Management Learning & Education **4**(2), 193–212 (Jun 2005). https://doi.org/10.5465/amle.2005.17268566, http://journals.aom.org/doi/10.5465/amle.2005.17268566
6. Park, J.S., O'Brien, J., Cai, C.J., Morris, M.R., Liang, P., Bernstein, M.S.: Generative Agents: Interactive Simulacra of Human Behavior. In: Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology. pp. 1–22. ACM, San Francisco CA USA (Oct 2023). https://doi.org/10.1145/3586183.3606763, https://dl.acm.org/doi/10.1145/3586183.3606763

7. Sottilare, R.A., Schwarz, J.: Adaptive Instructional Systems. Design and Evaluation: Third International Conference, AIS 2021, Held as Part of the 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part I, vol. 12792. Springer Nature (2021), https://books.google.com/books?hl=en&lr=&id=cbI2EAAAQBAJ&oi=fnd&pg=PR5&dq=R.+A.+Sottilare+and+J.+Schwarz+(Eds.),+Adaptive+Instructional+Systems:+Design+and+Evaluation,+Proc.+AIS+2021,+LNCS+12792,+Springer+Nature,+2021.+https://doi.org/10.1007/978-3-030-77857-6+&ots=SeVHstPSOY&sig=zvxLSVBwX0kzzIP5IVZSEieOEcg
8. The Design-Based Research Collective: Design-Based Research: An Emerging Paradigm for Educational Inquiry. Educational Researcher **32**(1), 5–8 (Jan 2003). https://doi.org/10.3102/0013189X032001005, https://journals.sagepub.com/doi/10.3102/0013189X032001005
9. Vaithilingam, P., Zhang, T., Glassman, E.L.: Expectation vs. Experience: Evaluating the Usability of Code Generation Tools Powered by Large Language Models. In: CHI Conference on Human Factors in Computing Systems Extended Abstracts. pp. 1–7. ACM, New Orleans LA USA (Apr 2022). https://doi.org/10.1145/3491101.3519665, https://dl.acm.org/doi/10.1145/3491101.3519665
10. Wagner, M.: FROM VIBE-TEACHING TO FLOW-TEACHING: REGROUNDING AI-ENHANCED PEDAGOGY IN GOAL-DIRECTED ENGAGEMENT. In: ICERI2025 Proceedings. pp. 769–778. IATED (2025), https://library.iated.org/view/WAGNER2025FRO2
11. Woo, D.J., Guo, K., Yu, Y.: A vibe coding learning design to enhance EFL students' talking to, through, and about AI (Sep 2025). https://doi.org/10.48550/arXiv.2509.08854, http://arxiv.org/abs/2509.08854, arXiv:2509.08854 [cs]
12. Zhu, X., Magee, L., Mischler, P.: Integrating Generative AI into LMS: Reshaping Learning and Instructional Design (Oct 2025). https://doi.org/10.48550/arXiv.2510.18026, http://arxiv.org/abs/2510.18026, arXiv:2510.18026 [cs]