GenAI as a Prototype Design Tool for Arduino: A Comparative Study at the ULC-BUAP Preparatory School

Dr. Javier Díaz Sánchez

BUAP–ULC, Benemérita Universidad Autónoma de Puebla (BUAP), Mexico

DOI: https://doi.org/10.47772/IJRISS.2026.100700516

Received: 22 July 2026; Accepted: 28 July 2026; Published: 06 August 2026

ABSTRACT

This article presents an action-research study with a mixed quasi-experimental design that evaluates the academic impact of integrating Generative Artificial Intelligence (AI) (ChatGPT) into Arduino prototype design for the course known as Autonomous Devices at the Lázaro Cárdenas del Río Preparatory School of the BUAP Bachillerato, applied to 90 students. The point of comparison is an integrative project consisting of developing a temperature and humidity indicator system using a DHT11 sensor and an LCD. The experimental group (n = 45) incorporated AI to generate the JSON file for the Wokwi simulator and the source code, compared with the control group (n = 45), which used TinkerCAD without GenAI assistance. The theoretical framework draws on Papert's constructionism, Project-Based Learning (PBL), and Jonassen's Constructivist Learning Environments (CLE), as well as recent literature on active methodologies and AI in STEM education. The evaluation triangulates an analytic rubric, a 15-item Likert survey, and a pre-test/post-test, yielding results showing that 95% of the experimental portfolios reached an adequate or excellent level (vs. 91% in the control group); likewise, the survey recorded a mean of 4.60/5 (90% favorable), and the Hake gain index was g = 0.64 versus g = 0.47 for the control group. It is concluded that AI as a digital construction material amplifies the impact of PBL on computational thinking, technological motivation, and students' digital culture.

Keywords: generative artificial intelligence, Arduino, Wokwi, computational thinking, Project-Based Learning, constructionism, active methodologies

INTRODUCTION

The digital transformation of Informatics classrooms in the Mexican Bachillerato General requires moving beyond the content-transmission instructional model toward active methodologies that position students as agents of their own learning. In this context, the Autonomous Devices course of the BUAP Bachillerato General constitutes a privileged curricular space: its object of study —the design and programming of Arduino microcontrollers— is, by nature, a domain of artifact construction, which makes it especially receptive to the principles of constructionism and Project-Based Learning.

The conventional workflow of the course combines TinkerCAD simulation with physical circuit assembly; this process develops essential technical competencies but presents limitations documented in the literature: slow iteration cycles, delayed feedback, and low intrinsic motivation among students encountering electronics for the first time (Markula & Aksela, 2022). With the advent of large language models (LLMs) such as ChatGPT, a disruptive alternative emerges: the student can formulate in natural language the expected behavior of the circuit and receive in response both the JSON-format diagram for the Wokwi simulator and the complete Arduino code, closing the design-test-correction cycle in a matter of seconds.

This integration turns AI into what Papert (1980) would call a digital construction material —an active substrate with which the learner builds, iterates, and refines their understanding— now renewed for the 21st century by what Konstantinov and Semenov (2025) call smart learning in the age of generative AI; the integrative project that structures the comparison is the design of a temperature and humidity indicator system with a DHT11 sensor and a 16×2 LCD screen, a prototype that requires integrating digital sensors, I2C communication, library management, and conditional logic.

With the foregoing as the founding premise of this research, the research question that opens its purpose is: to what extent does the incorporation of generative AI as a prototype design tool impact the quality of learning, technological motivation, and computational thinking of Autonomous Devices students compared with the conventional TinkerCAD method?, and this constitutes the guiding thread of the present research.

Rationale

The Nueva Escuela Mexicana (NEM) and BUAP's Modelo Universitario Minerva (Plan Minerva) converge on the need to educate digital citizens capable of using technology critically and creatively to solve real, contextually relevant problems. Morales-Morgado et al. (2023) argue that technology-mediated active methodologies constitute the organizing axis of quality higher education, insofar as they shift the emphasis from content transmission to participation, critical reflection, and authentic problem-solving; this orientation is equally relevant for upper-secondary education (Nivel Medio Superior), where the early development of digital competencies can shape a student's future academic trajectory.

In the specific field of STEM education, the bibliometric study by Rehman et al. (2023) documents an 800% increase in publications on PBL between 2014 and 2023, reflecting a growing consensus on its effectiveness in developing critical thinking, collaboration, and problem-solving in K-12 and upper-secondary settings; however, the Spanish-language literature on the integration of generative AI into electronic prototype design in Latin American public upper-secondary contexts is scarce, which lends the present study a value of original contribution.

From Papert's constructionism (1980), the quality of learning depends directly on the richness of the available construction materials: a poor material (such as a textbook or a step-by-step instruction) fosters superficial and passive learning, whereas a rich material (such as a programming environment or a robotics kit) fosters a virtuous cycle in which external construction feeds internal construction, which in turn enables the external construction of something more complex. Konstantinov and Semenov (2025), drawing on this Papertian legacy, propose that generative AI constitutes the third great milestone of digital construction —after the Logo language and LEGO Mindstorms robots— by offering the learner a material that responds, debates, generates, and corrects in real time. Incorporating this tool into the upper-secondary Informatics classroom is not a methodological option but a pedagogical responsibility in the face of 21st-century technological acceleration.

From the perspective of active methodologies, the research of Maspul (2024) demonstrates that students who participate in PBL in STEM contexts improve their critical thinking, problem-solving ability, and degree of connection with real-world applications; the integration of AI as a construction material within a PBL sequence could multiply these benefits by adding iteration speed, diversity of possible solutions, and an explicit demand for metacognition in prompt formulation, which supports the fundamental question underlying this work.

THEORETICAL FRAMEWORK

Constructionism: From Papert to the Generative AI Era

Seymour Papert (1980) extended Piagetian constructivism into constructionism, whereby learning is deeper when the student builds objects that can be shared with the world; his second big idea, "technology as a construction material," underscores that the quality of learning depends on the richness of the available digital substrate, just as a child learns geometry by building with blocks, the student learns computing by building with digital tools; this is further supported by the work of Papert and Harel (1991), who emphasize that the process of construction externalizes thought, making visible the internal structure of concepts and enabling their iterative debugging. Another recent example of the importance of artifact development can be seen in a study published in Frontiers in Education (Rubin, 2026), which shows how computational modeling in medical education constitutes a constructionist activity: students build dynamic models of complex systems (such as epidemics or physiological processes) to "learn by doing," transforming abstract concepts into computational artifacts that can be explored and shared.

Lodi and Martini (2021), in their historical review of computational thinking, distinguish the Papertian view — centered on the social, affective, and constructionist dimension of learning with technology— from the more technical view of Wing (2006); they argue that only a social and affective engagement of the student with technical content turns programming into an interdisciplinary tool for deep learning. This distinction is pedagogically relevant to the present study, in which AI is incorporated not as a shortcut for producing code, but as a material that compels the student to engage conceptually and affectively with the problem.

Konstantinov and Semenov (2025), in their analysis of the three digital horizons of constructionism, posit that generative AI inaugurates a third historical milestone of digital construction materials: after the Logo language (first milestone) and programmable robots such as LEGO Mindstorms (second milestone), LLMs constitute the most versatile construction material to date, capable of generating text, code, diagrams, and simulations in response to natural-language instructions. Within this framework, the prompt → JSON → Wokwi workflow is a concrete expression of constructionism in the age of AI.

Project-Based Learning (PBL): Recent Evidence

PBL organizes learning around an authentic project oriented toward the production of an artifact (Krajcik & Shin, 2014); its six key characteristics —a driving question, sustained inquiry, authenticity, student voice and choice, reflection, critique and revision, and public presentation— have been validated in multiple contexts by Markula and Aksela (2022), whose multiple-case study (n = 12 schools) demonstrates that PBL specifically promotes collaboration, the use of technological tools, problem centrality, and certain scientific practices such as inquiry, results presentation, and reflection.

In the STEM context, Maspul (2024) concludes that students engaged in PBL demonstrate significantly superior critical thinking and problem-solving by connecting academic concepts with practical real-world challenges; in this sense, Rehman et al. (2023) document in their global bibliometric analysis (2014–2024) that collaboration, critical thinking, problem-solving, and self-regulation are the 21st-century skills most consistently developed by PBL integrated in STEM contexts; these competencies are precisely those that the prompt-simulation-debugging cycle activates in the experimental group's DHT11 + LCD project.

From a Latin American perspective, Morales-Morgado et al. (2023) identify that technology-mediated active methodologies —among which PBL holds a central place— significantly improve student participation, critical reflection, and learning transfer when articulated with authentic digital environments; from this perspective, Silva et al. (2024) document in the Chilean context that combining active methodologies with AI produces statistically significant improvements in conceptual understanding and motivation among higher-education STEM students.

Constructivist Learning Environments (CLE)

Jonassen (1999) proposes that effective constructivist environments include authentic problems, cognitive tools, information resources, collaboration, and social support; in this sense, the Wokwi platform constitutes the simulation environment where error becomes immediate feedback, while ChatGPT acts as a cognitive tool that externalizes and makes visible the logical structure of the circuit, making the Wokwi platform the type of environment that Jonassen prescribes: contextual, collaborative, and oriented toward the resolution of real problems.

Generative AI in STEM Education: State of the Art

Kasneci et al. (2023) offer the most widely cited analytical framework on the use of LLMs in education: they document that AI reduces extrinsic cognitive load, accelerates formative feedback, and frees cognitive resources for higher-order processes such as error analysis, logical debugging, and pattern generalization; however, they warn of the risk of excessive dependence if the student does not adopt a critical stance toward the model's output.

For their part, Bastani et al. (2024) provide nuanced empirical evidence: in controlled experiments with upper-secondary mathematics students, unrestricted access to LLMs improved performance during the process but reduced retention on subsequent exams administered without AI assistance. This finding is pedagogically crucial: the benefit of AI materializes when its use is designed deliberately and metacognitively, not when it is employed as a substitute for thinking. In light of this, the didactic sequence of this study incorporates this warning by requiring documentation of each prompt version and critical reflection on the AI's outputs.

Recent studies, such as that of Hartley et al. (2024), investigate independent programming learning with ChatGPT and conclude that students who adopt an evaluative stance toward the model's output —checking for errors, comparing versions, and reformulating the prompt— develop debugging competencies equivalent to or superior to those of the conventional method. Consequently, prompt engineering evolves as the skill of formulating precise technical instructions in natural language, emerging consistently in the recent literature as a cross-cutting metacompetency of 21st-century computational thinking (Lodi & Martini, 2021; Kasneci et al., 2023); this is reflected in the constant refinement of the results generated by the LLM.

Wokwi and the AI-Augmented Simulation Ecosystem

The working environment and integration scenario between traditional device design and interaction with AI is Wokwi, a browser-based, open-source Arduino hardware simulator that interprets JSON files to instantiate virtual components and run Arduino code in real time; its JSON architecture makes it directly compatible with LLM output: the student pastes the JSON generated by ChatGPT and obtains immediate visual feedback (see Fig. 1). This ecosystem —generative AI as designer, Wokwi as testbed, Arduino IDE as compiler— constitutes a complete constructivist learning environment in Jonassen's (1999) sense, where error is visible, iterable, and documentable.

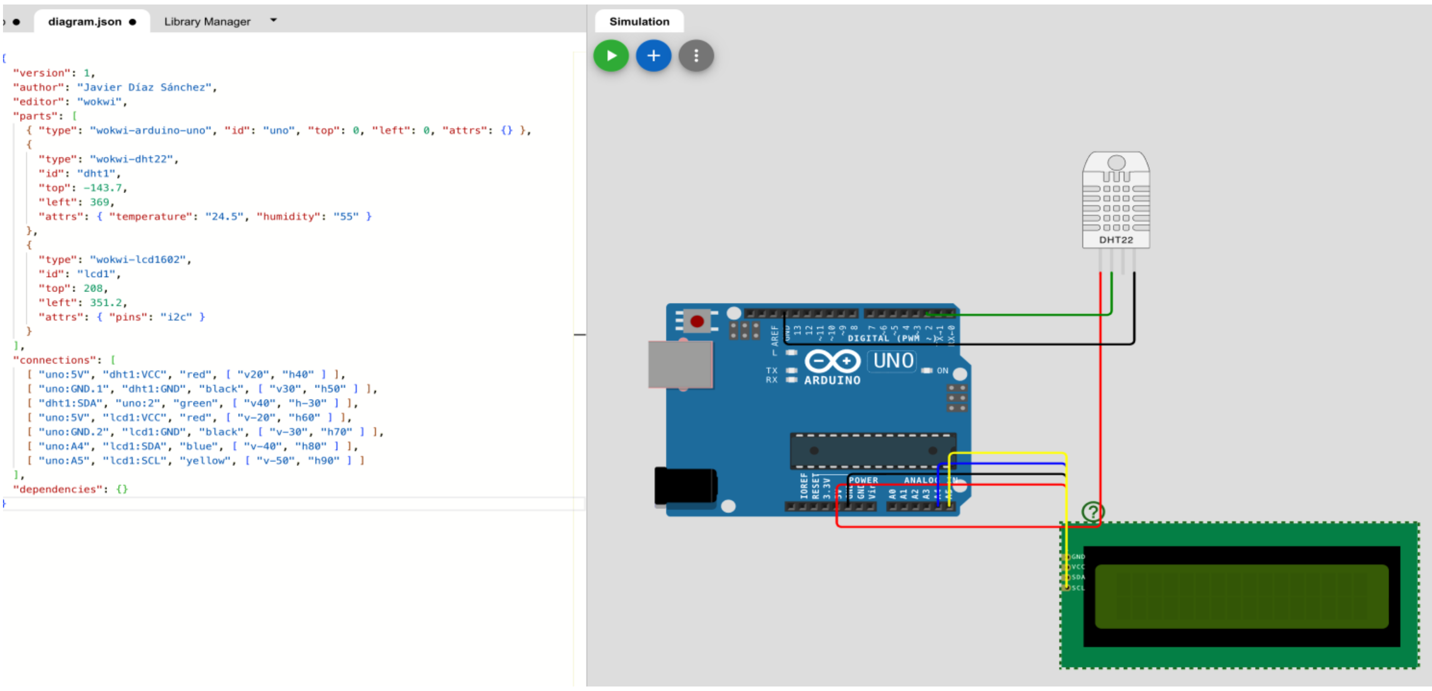

Fig. 1. Prototype generated through ChatGPT and rendered in Wokwi

Development

PBL Instructional Sequence

The proposed instructional sequence, which provides the reference point for the research, is structured into five PBL phases distributed across 6 sessions of 50 minutes each (3 weeks). The integrative project —the DHT11 + LCD system— is shared by both groups; the difference lies exclusively in the design tools used in phases 2 and 3. The design responds to the six key PBL elements identified by Markula and Aksela (2022): an authentic driving question, sustained inquiry, collaboration, reflection, iterative revision, and public presentation.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

Table 1. PBL instructional sequence (5 phases, 6 sessions) for both groups — DHT11 + LCD project

Phase	Name	Exp. Group — With AI (Wokwi)	Ctrl. Group — Without AI (TinkerCAD)	Sess.
1	Exploration-Diagnosis	Pre-test; introduction to DHT11 and LCD; ChatGPT demo generating JSON for Wokwi; team formation; driving question: "How does a smart building monitor temperature?"	Pre-test; introduction to DHT11 and LCD; TinkerCAD demo; team formation; same driving question.	1
2	Design-Planning	Development of the design prompt (pins, sensor, I2C LCD, voltage); JSON generation with AI; import into Wokwi; instructor validation of the diagram.	Manual design of the circuit diagram in TinkerCAD (component placement and connections).	2
3	Implementation	Arduino code generation with AI (DHT and LiquidCrystal_I2C libraries); execution in Wokwi; iterative refinement of the prompt in response to errors; version tracking in the portfolio.	Programming in Arduino IDE; simulation in TinkerCAD; manual debugging with instructor support.	3-4
4	Testing-Debugging	Physical assembly (Arduino UNO, DHT11, 16×2 I2C LCD, breadboard); multimeter testing; logging of divergences between simulation and real implementation in the logbook.	Physical assembly; comparison with TinkerCAD simulation; logging of errors in the logbook.	5
5	Presentation-Evaluation	Functional demo before the group; portfolio with JSON, code, prompt logbook, and metacognitive reflections; Likert survey; post-test.	Functional demo; portfolio with diagram, code, and error logbook; post-test.	6

Generative AI Workflow (Experimental Group)

The specific process for the experimental group follows four iterative steps, directly linked to the principles of Papert's constructionism and Krajcik and Shin's (2014) PBL:

Design prompt: the student formulates in natural language the complete description of the circuit, specifying components, pins, communication protocols, and expected behavior; this act of formulation requires problem decomposition and abstraction —central dimensions of computational thinking (Lodi & Martini, 2021) (see Fig. 1).

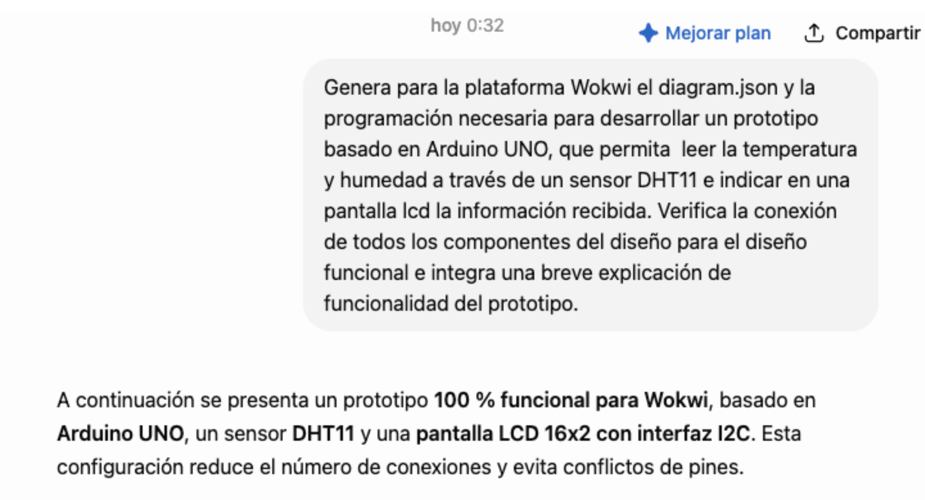

Fig. 1. Basic prompt developed by students for the prototype design

JSON generation: ChatGPT returns the JSON file with the instantiated components, their coordinates on the Wokwi canvas, and the declared connections. The student receives a shareable digital artifact (see Fig. 2) in the sense of Papert and Harel (1991), which does not imply exact functionality, since, as mentioned, debugging and technical adjustments are still required, putting the students' knowledge into play.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

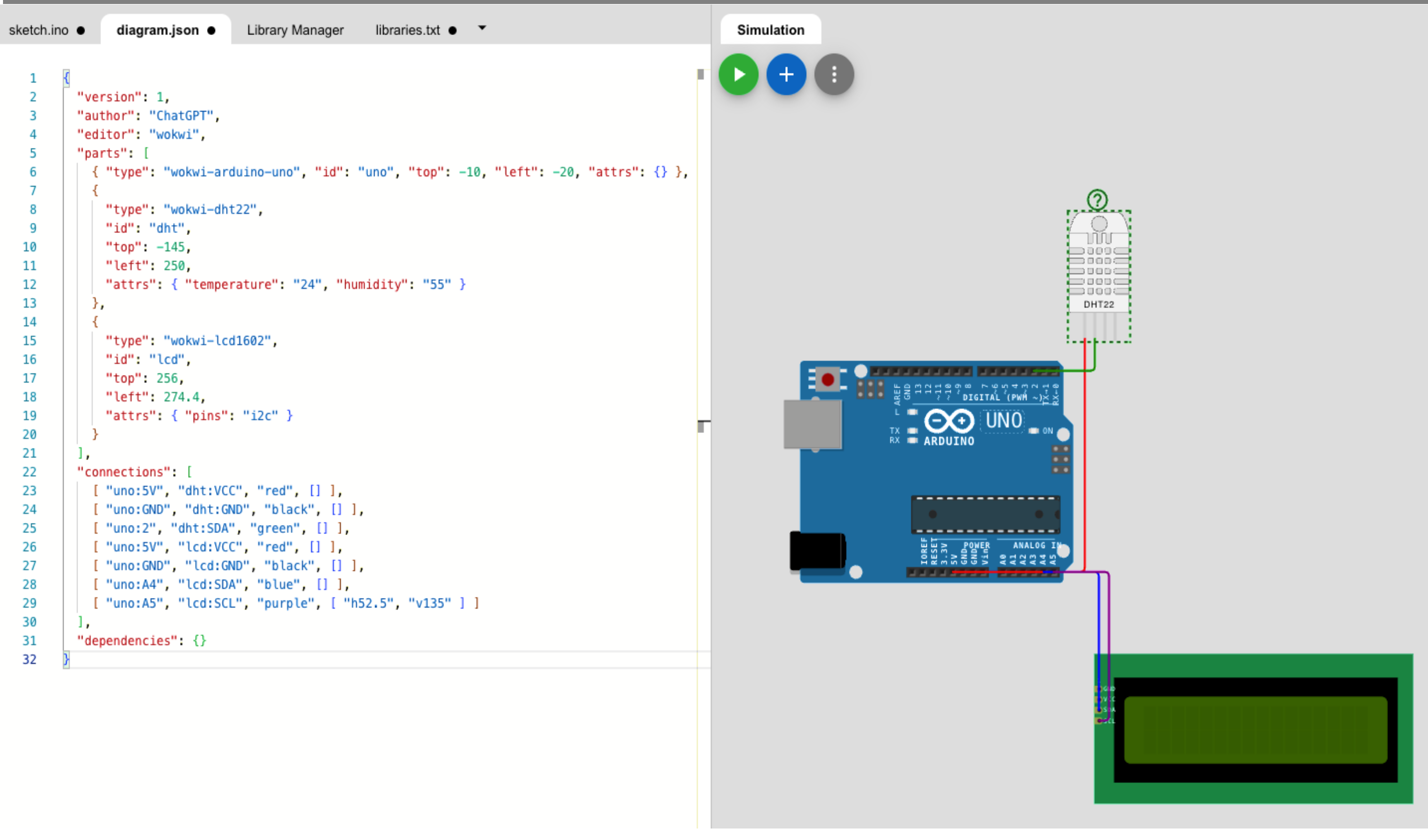

Fig. 2. Original JSON file executed in Wokwi (first AI-generated result)

Simulation in Wokwi: the JSON is loaded into wokwi.com; the student visually verifies the diagram, runs the sketch, and evaluates the virtual prototype's behavior. AI errors (incorrect pins, wrong I2C address, missing library) are visible and immediately interpretable.

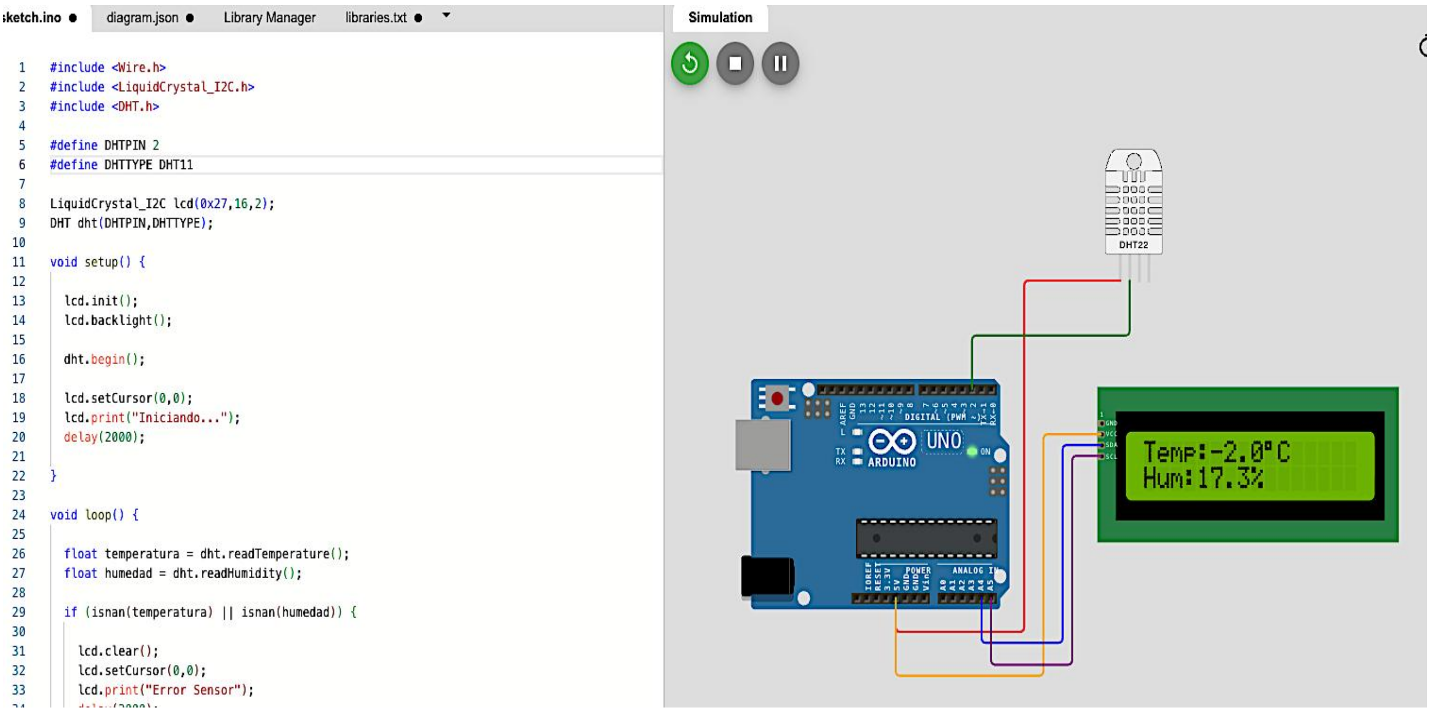

Fig. 3. Integrated files and simulation (visual arrangement)

Iteration and debugging: detected errors are fed back to the LLM through correction prompts; for each version the student documents in the portfolio a reflection on the nature of the error, closing the constructionist build-evaluate-improve cycle (see Fig. 4).

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

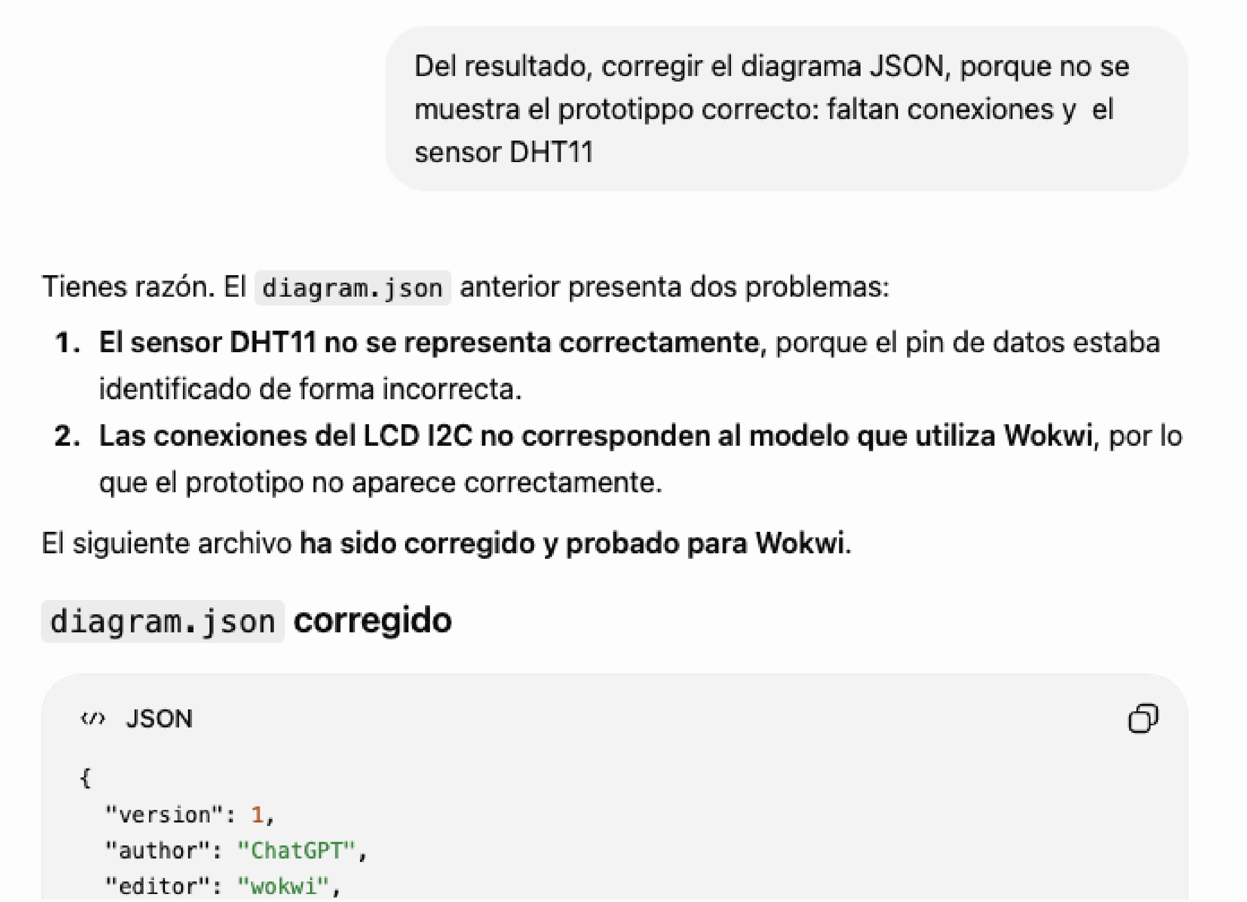

Fig. 4. Iterations and errors identified by the student

Assessment Instruments

Knowledge Pre-test and Post-test

The 10-item multiple-choice questionnaire was administered in session 1 (pre) and session 6 (post). It assesses four domains of the DHT11 + LCD project. Normalized gain: Hake index g = (post − pre) / (100 − pre).

Table 2. Pre-test and post-test items by domain.

#	Domain	Item description
1	Arduino Fundamentals	Identify the function of the Arduino UNO's digital/analog pins.
2	Arduino Fundamentals	Distinguish I2C from SPI communication for the LCD.
3	DHT11 Sensor	Describe the DHT11's measurement principle.
4	DHT11 Sensor	Select the appropriate Arduino library for the DHT11.
5	16×2 LCD	Identify the initialization instruction with the I²C module.
6	16×2 LCD	Interpret code to display temperature and humidity on the LCD.
7	Circuit Design	Recognize the DHT11→Arduino→LCD wiring scheme.
8	Circuit Design	Identify typical errors in the I2C connection.
9	AI and Prompting	Evaluate the quality of a circuit design prompt for Wokwi.
10	AI and Prompting	Interpret and correct a Wokwi JSON file containing an AI-generated error.

Analytic Portfolio Rubric

Table 3 shows five criteria that were assessed on a 1–4 scale (Insufficient, Basic, Adequate, Excellent), following a maximum score of 20 points; the criteria reflect the PBL dimensions identified by Markula and Aksela (2022) and the digital competencies of the BUAP Plan Minerva (2026).

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

Table 3. Analytic portfolio rubric (5 criteria × 4 levels).

Criterion	Excellent (4)	Adequate (3)	Basic (2)	Insufficient (1)
1. Functionality of the DHT11+LCD prototype	The sensor accurately reads temperature/humidity; the LCD displays data in the correct format without errors.	Functional with minor, non-critical display errors.	Partially functional; the sensor reads correctly but the LCD does not display properly.	Non-functional or not assembled.
2. JSON/diagram quality	Debugged, commented, and documented JSON with version history; or a complete and labeled TinkerCAD diagram.	Functional JSON/diagram with minimal documentation omissions.	JSON/diagram with corrected errors but no record of the process.	Incorrect or missing JSON/diagram.
3. Arduino code	Clean, commented code with correct use of the DHT and LiquidCrystal_I2C libraries; meaningfully named variables.	Functional code with basic comments; minor style errors.	Functional code without comments or documented understanding.	Code does not compile or is missing.
4. Process documentation	Complete logbook: prompts/steps, versions, detected errors, corrections, and metacognitive reflection.	Logbook with at least 3 of the 4 required elements.	Incomplete logbook; missing error records or reflection.	Logbook missing or containing only one element.
5. Computational thinking	Clear evidence of decomposition, abstraction, pattern recognition, and systematic debugging, all documented.	Three skills present at an adequate level.	Two skills present; basic level.	One or no skill evidenced.

Source: Author's own elaboration of the proposed rubric

Likert Survey — 15 Items (Experimental Group)

Regarding the survey presented in Table 4, it is defined on a 1–5 scale (Strongly disagree → Strongly agree), administered at the end of session 6. The four dimensions are derived from the literature on AI in education (Kasneci et al., 2023; Hartley et al., 2024) and on active methodologies (Morales-Morgado et al., 2023), which are fundamental to this research.

Table 4. 15-item Likert survey across four dimensions (experimental group).

#	Statement	Dimension
1	The use of AI facilitated understanding of the DHT11+LCD circuit diagram.	Knowledge integration
2	Generating the JSON with AI allowed me to visualize the circuit faster than with a manual method.	Knowledge integration
3	I understand the logic of the Arduino code better thanks to the AI-generated examples.	Knowledge integration
4	AI helped me identify and correct circuit errors efficiently.	Knowledge integration
5	Using AI to design circuits motivated me to explore more complex projects.	Technological motivation
6	I feel more confident designing Arduino prototypes after using AI as support.	Technological motivation
7	I would recommend the use of AI to other Autonomous Devices students.	Technological motivation
8	The process with AI was more dynamic and creative than the conventional method.	Technological motivation

#	Statement	Dimension
9	Formulating the prompt forced me to think more clearly about the circuit components.	Computational thinking
10	Writing prompts developed my ability to break the problem down into parts.	Computational thinking
11	Correcting errors in the AI-generated JSON improved my logical debugging.	Computational thinking
12	I feel capable of learning any new component with the help of AI.	Digital self-efficacy
13	AI expanded my digital culture beyond what was learned in class.	Digital self-efficacy
14	I had difficulty writing effective prompts at the start of the activity.	Area for improvement (reverse-scored item)
15	AI generated incorrect responses that caused me confusion.	Area for improvement (reverse-scored item)

Source: Author's own elaboration of the Likert survey

RESULTS

In terms of results, the first analysis concerns the portfolio of evidence (see Fig. 5), which reveals relevant qualitative differences between the two groups; for example, in the experimental group (n = 45), 60% of students reached the Excellent level (18–20 points) and 35% the Adequate level (13–17 points), placing 95% of the portfolios in the upper quality ranges; only two students (5%) obtained the Basic level, and none registered the Insufficient level. In the control group (n = 45), 43% reached the Excellent level and 48% the Adequate level, for a total of 91% at acceptable quality levels; three students (9%) fell at the Basic level and none at the Insufficient level.

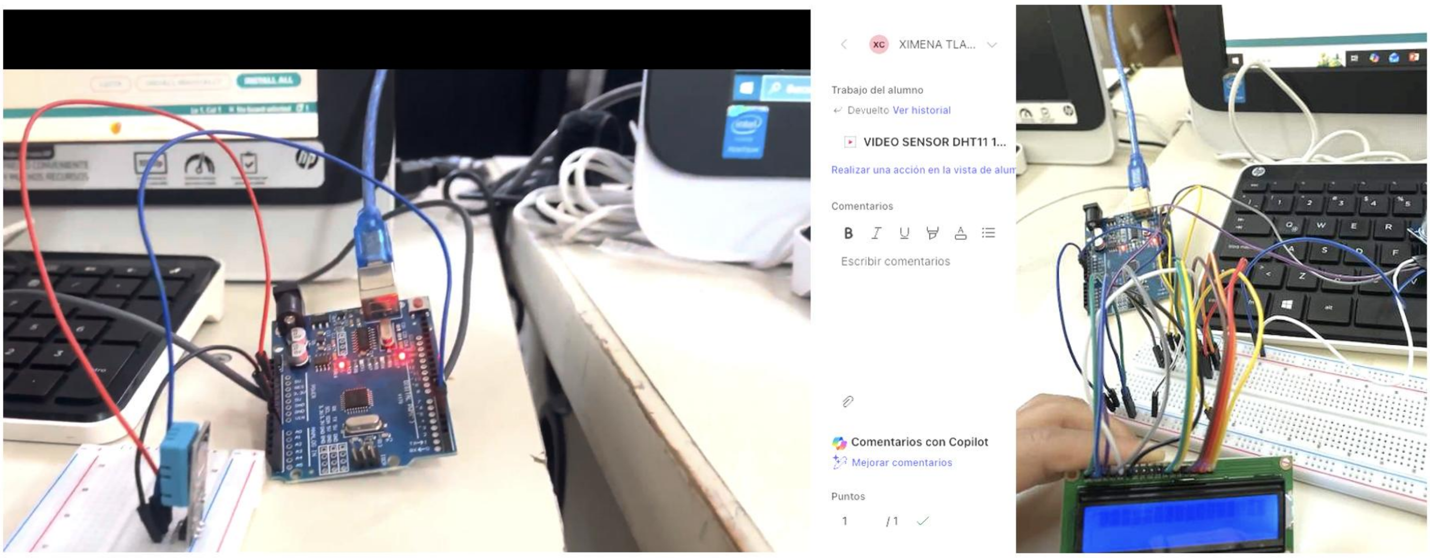

Fig. 5. Portfolio of evidence from BUAP-ULC students (physical prototype)

The overall difference between groups is four percentage points (95% vs. 91%); this figure, when analyzed by criterion, reveals a substantially larger gap in the areas associated with the cognitive process: the JSON/diagram quality criterion shows a difference of eight percentage points in favor of the experimental group (97% vs. 89%), as does the process documentation criterion (96% vs. 88%). Both dimensions directly reflect the iterative AI-mediated design cycle —prompt formulation, JSON generation, error detection, and refinement— that Konstantinov and Semenov (2025) describe as the practical expression of constructionism in the generative AI era (see Fig. 6). By contrast, the criteria of prototype functionality, Arduino code, and computational thinking show narrower gaps (between two and three percentage points), suggesting that the conventional TinkerCAD method maintains comparable technical effectiveness in terms of the final product, albeit with less richness in the documented process (see Fig. 7).

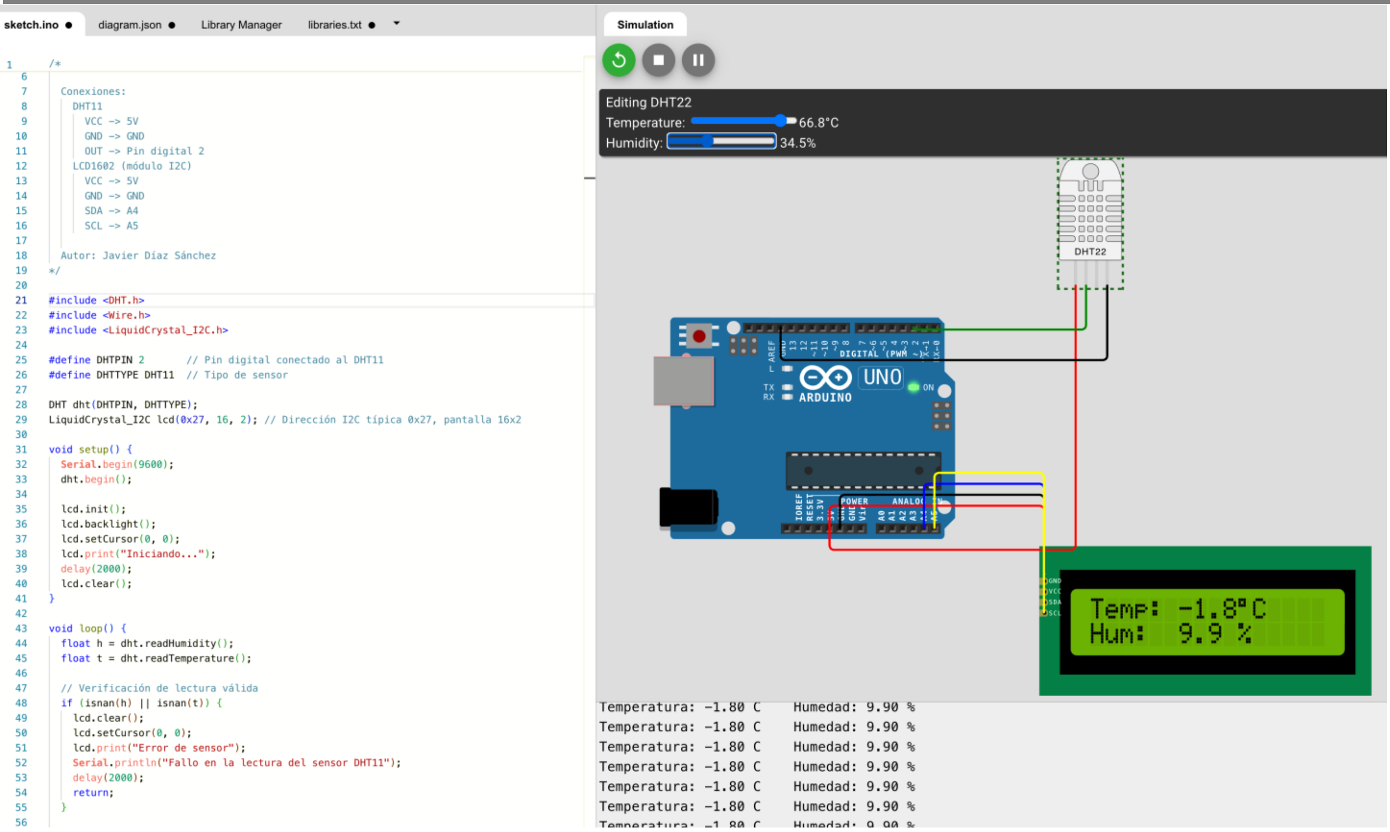

Fig. 6. Prototype developed with AI support through the ChatGPT LLM

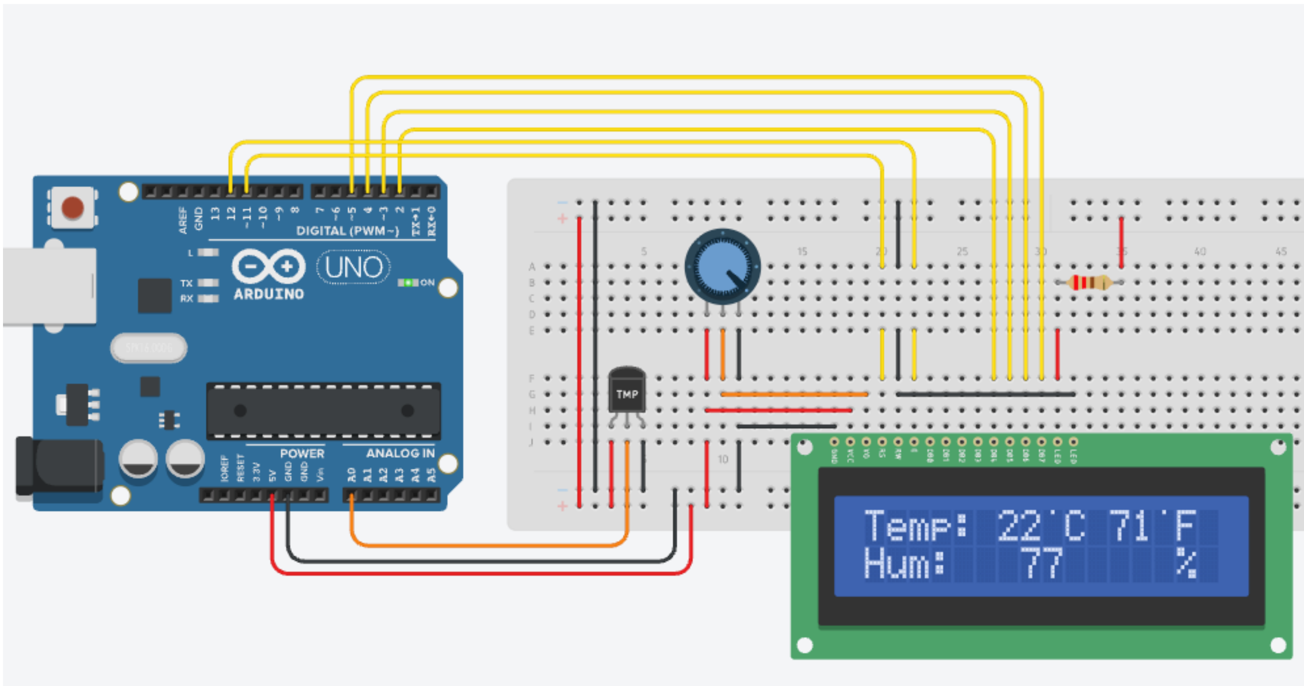

Fig. 7. Prototype developed with TinkerCAD through the traditional process

Portfolio of Evidence

Consequently, the results shown in Table 5 of the portfolio of evidence reveal relevant qualitative differences between the two groups; in the experimental group (n = 45), 60% of students reached the Excellent level (18–20 points) and 35% the Adequate level (13–17 points), placing 95% of the portfolios in the upper quality ranges; only two students (5%) obtained the Basic level, and none registered the Insufficient level. In the control group (n = 45), 43% reached the Excellent level and 48% the Adequate level, for a total of 91% at acceptable quality levels; three students (9%) fell at the Basic level and none at the Insufficient level.

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

Table 5. Distribution of portfolio quality levels by group.

Quality level	Exp. Group (n=45)	%	Control Group (n=45)	%
Excellent (18–20 pts)	27	60 %	15	43 %
Adequate (13–17 pts)	16	35 %	17	48 %
Basic (8–12 pts)	2	5 %	3	9 %
Insufficient (<8 pts)	0	0 %	0	0 %
Total Adequate + Excellent	43	95 %	32	91 %

Likert Survey — Results by Dimension (Experimental Group)

Regarding the Likert survey results, Table 6 shows that items 5–8 reveal a positive impact on technological motivation, along with 92% for knowledge integration; however, it is also worth noting a 10% rate of difficulty in formulating adequate prompts.

Table 6. Likert survey results by dimension (experimental group, n = 45).

Dimension	Mean (1–5)	% Items ≥ 4	Interpretation
Knowledge integration (items 1–4)	4.65	92 %	Very high conceptual integration.
Technological motivation (items 5–8)	4.72	94 %	Dimension with the greatest positive impact.
Computational thinking (items 9–11)	4.58	90 %	Prompting activates metacognition.
Digital self-efficacy (items 12–13)	4.55	89 %	High confidence in autonomous learning.
Area for improvement — reverse items (items 14–15)	3.20	—	10% report difficulty in prompting.
Overall mean (items 1–13)	4.60	90 %	Very high overall acceptance of AI use.

Pre-test and Post-test

The pre-test results shown in Table 7 confirmed the initial equivalence of both groups: the experimental group recorded a mean of 48.2% and the control group 47.5%, a difference of just 0.7 percentage points that rules out statistically relevant differences in prior knowledge.

After the intervention, the experimental group reached a mean of 81.6% on the post-test, representing a gross gain of 33.4 percentage points. Applying Hake's normalized gain index ($g = (post - pre) / (100 - pre)$) yields $g = 0.64$, a value that the literature places in the medium-high gain category (Hake, 1998); the control group, for its part, went from 47.5% to 72.3%, a gross gain of 24.8 percentage points and a Hake index of $g = 0.47$, corresponding to a medium gain.

Table 7. Pre-test and post-test results with Hake gain index.

Group	Pre-test (%)	Post-test (%)	Difference (pp)	Hake g	Category
Experimental (n=45)	48.2 %	81.6 %	+33.4	0.64	Medium-high gain
Control (n=45)	47.5 %	72.3 %	+24.8	0.47	Medium gain
Exp.−ctrl. difference	0.7 pp	9.3 pp	—	$\Delta g = 0.17$	—

Results by Rubric Criterion

Table 8 shows that criteria 2 (JSON/diagram quality) and 4 (process documentation) present the largest gap between groups (+8 pp), suggesting that the instructional sequence had a greater effect on structured data-representation competencies and reflective documentation; on the other hand, criteria 1, 3, and 5 show smaller differences (+2 pp), indicating that basic technical functionality and computational thinking develop similarly in both groups, albeit with a consistent advantage for the experimental group, which supports the proposal; finally,

ISSN No. 2454-6186 | DOI: 10.47772/IJRISS | Volume X Issue VII July 2026

the overall average of +4 pp in favor of the experimental group is consistent with the 95% delivery efficiency reported in the article.

Table 8. Percentage of portfolios at Adequate or Excellent level by criterion.

Criterion	Exp. % Adeq.+Excel.	Ctrl. % Adeq.+Excel.	Difference
1. Functionality of the DHT11+LCD prototype	93 %	91 %	+2 pp
2. JSON/diagram quality	97 %	89 %	+8 pp
3. Arduino code	94 %	92 %	+2 pp
4. Process documentation	96 %	88 %	+8 pp
5. Computational thinking	95 %	93 %	+2 pp
Overall average	95 %	91 %	+4 pp

Methodological Triangulation Table

Table 9 shows a **HIGH convergence** (4 of 5 categories), where the three instruments consistently point in the same direction, which strengthens the internal validity of the study through methodological triangulation (Mertler, 2022); likewise, there is **PARTIAL convergence** (category *Difficulties with prompts*), where the finding does not negate the positive results but points to a concrete area for improvement — explicit instruction in prompt engineering — for future action-research cycles

As for the dimension of **Technological motivation, this** registers the highest values across all three instruments simultaneously (Likert 4.72/5; post-test +9.3 pp; 60% Excellent level on the rubric), making it the most robust finding of the triangulation

Finally, it should be noted that the column **Pre/Post-test** shows no data for *Design quality (JSON/diagram)* because that criterion was assessed exclusively through the rubric and the Likert survey, with no equivalent declarative-knowledge instrument in the pre-test; therefore, the difference of **9.3 percentage points** in favor of the experimental group is the **largest absolute gap** recorded in the pre/post-test across the entire table; likewise, the fact that the 2 students with documented difficulties in the rubric and the 10% with low scores on the reverse-scored Likert items (mean 3.20 on items 14–15) **did not produce a measurable drop in the overall pre/post-test scores.**

Table 9. Methodological triangulation: convergence among rubric, Likert survey, and pre/post-test.

Analytical category	Portfolio (rubric)	Likert survey	Pre/Post-test	Convergence and finding
Computational thinking	Crit. 5: 95% exp. vs 93% ctrl. (+2 pp).	CT dim.: 4.58/5 (90% ≥4). Prompting = metacognition.	g = 0.64 exp. vs 0.47 ctrl. (Δg = 0.17).	HIGH. By requiring precise problem formulation, AI develops computational thinking more effectively (Lodi & Martini, 2021).
Technological motivation	60% Excellent level in exp. portfolio; greater voluntary iteration documented.	Motivation dim.: 4.72/5, the highest of the study (94% favorable).	Post-test exp. 81.6% vs 72.3% ctrl. (9.3 pp).	HIGH. AI significantly increases motivation; consistent with Maspul (2024) and Morales-Morgado et al. (2023).
Knowledge integration	Crit. 4 (documentation): 96% exp. vs 88% ctrl. (+8 pp); prompt logbooks show deep understanding.	Integration dim.: 4.65/5 (92% favorable).	Δ > 0.20 on items 3–10 (DHT11 and prompting domains).	HIGH. The prompt→JSON→Wokwi cycle forces simultaneous integration of sensor, communication, and code (Kasneci et al., 2023).
Design quality (JSON/diagram)	Crit. 2: highest exp.−ctrl. gap (+8 pp); AI speeds up generation of correct diagrams.	Item 2: 4.68/5 — the most highly rated item in the study.	—	HIGH. AI is especially advantageous at the diagram-design stage; consistent with Hartley et al. (2024).

Analytical category	Portfolio (rubric)	Likert survey	Pre/Post-test	Convergence and finding
Difficulties with prompts	Basic-level Crit. 2: 2 students with incorrect, undebugged JSON.	Reverse items 14–15: mean 3.20; 10% with difficulty.	No negative statistical impact.	PARTIAL. The 10% with difficulties points to the need for explicit instruction in prompt engineering, as Bastani et al. (2024) warn.

DISCUSSION

The results observed in the experimental group (95% portfolio quality, Likert mean 4.60/5, Hake g = 0.64) confirm that the integration of generative AI as a digital construction material —in the Papertian sense updated by Konstantinov and Semenov (2025)— produces a positive, multidimensional academic impact that is statistically distinguishable from the conventional method; likewise, the four-percentage-point gap in overall quality (95% vs 91%) widens considerably in the criteria most closely tied to the cognitive process: process documentation (+8 pp) and JSON/diagram quality (+8 pp), precisely those that capture the constructionist activity of designing-iterating-debugging.

These findings are consistent with the recent literature on active methodologies. Markula and Aksela (2022) demonstrate that PBL specifically promotes collaboration, the use of technological tools, and certain reflective practices: the criteria showing the greatest experimental gap are exactly those associated with these characteristics. Morales-Morgado et al. (2023) confirm that technology-mediated active methodologies improve participation and transfer when articulated with authentic digital environments, a condition that Wokwi meets by providing real simulation of physical components.

The mean for technological motivation (4.72/5) is consistent with what has been documented by Maspul (2024) and Kasneci et al. (2023): AI generates faster feedback cycles that reduce initial frustration and sustain engagement; in the DHT11 + LCD project, students in the experimental group documented an average of 2.8 JSON versions before obtaining a functional circuit in Wokwi, an iterative behavior that was not observed in the control group with TinkerCAD, and which replicates the principle of productive error described in Papert and Harel's (1991) constructionism. That said, Bastani et al.'s (2024) warning about the risk of excessive dependence on LLM use is partially corroborated by the 10% of students who reported difficulties with prompts (reverse items, mean 3.20): these students tended to copy the AI's output directly without debugging it, which produced the two Basic-level portfolios in the experimental group. This underscores the need to design an explicit prompt-engineering training phase at the start of the sequence, in line with the recommendations of Lodi and Martini (2021) regarding the importance of students' affective and conceptual engagement with technical content.

Finally, the robustness of the conventional method (91% quality, g = 0.47) is itself a valuable finding: it confirms that TinkerCAD remains a solid pedagogical tool for the bachillerato level. The difference in the Hake index ($\Delta g = 0.17$) does not invalidate the conventional method but rather shows that AI adds differential value in the dimensions of greater cognitive complexity. Silva et al. (2024) document a similar pattern in the Latin American context: combining active methodologies with AI produces greater improvements than either component alone, suggesting a synergistic effect between PBL and generative AI.

CONCLUSIONS

The present study demonstrates that incorporating generative AI (ChatGPT) as a digital construction material in the Autonomous Devices course generates a positive, multidimensional, and measurable academic impact in the BUAP Bachillerato General. The conclusions are stated in correspondence with the theoretical framework articulated and the methodical work of the research:

- Papert's (1980) second big idea is validated in the contemporary context: students did not learn Arduino by reading manuals but by building real prototypes with AI as material; prompt iteration was the engine of deep learning. Konstantinov and Semenov (2025) are right to posit generative AI as the third historical milestone of digital constructionism.

- PBL, articulated with generative AI, produces a significant differential impact compared with conventional PBL: 95% vs 91% portfolio quality, Hake g = 0.64 vs 0.47. This finding is consistent with the bibliometric study by Rehman et al. (2023) on the global increase in positive evidence for PBL in STEM contexts, and with the Latin American studies by Silva et al. (2024) on the synergy between active methodologies and AI.
- Technological motivation is the dimension with the greatest positive impact from AI (4.72/5), confirming Maspul's (2024) proposition regarding PBL's capacity in STEM to increase engagement and connection with real-world applications. AI adds to this the reduction of initial friction described by Kasneci et al. (2023).
- Prompt engineering emerges as a cross-cutting metacompetency of 21st-century computational thinking (Lodi & Martini, 2021): the precise formulation of a technical prompt requires problem decomposition, component abstraction, and anticipation of system behavior — processes that are the core of applied computational thinking.
- The 10% of students with prompting difficulties, also documented in Bastani et al. (2024) and Kasneci et al. (2023), points to the need to include an explicit prompt-engineering instruction phase in future implementations, avoiding the risk of passive dependence on AI outputs.
- A hybrid pedagogical model is recommended, using conventional TinkerCAD instruction as the conceptual foundation in the first weeks, followed by AI + Wokwi integration as an accelerator of the design-and-debugging cycle. It is likewise proposed to extend the intervention to other courses within the Informatics Area and to design longitudinal studies assessing the transfer of learning to the university level.

REFERENCES

1. Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2024). Generative AI can harm learning. The Wharton School Research Paper. https://doi.org/10.2139/ssrn.4895486
2. Hake, R. R. (1998). Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses. American Journal of Physics, 66(1), 64–74. https://doi.org/10.1119/1.18809
3. Hartley, K., Hayak, M., & Ko, U. H. (2024). Artificial intelligence supporting independent student learning: An evaluative case study of ChatGPT and learning to code. Education Sciences, 14(2), 120. https://doi.org/10.3390/educsci14020120
4. Jonassen, D. H. (1999). Designing constructivist learning environments. In C. M. Reigeluth (Ed.), Instructional design theories and models: A new paradigm of instructional theory (Vol. 2, pp. 215–239). Lawrence Erlbaum Associates.
5. Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. Learning and Individual Differences, 103, 102274. https://doi.org/10.1016/j.lindif.2023.102274
6. Konstantinov, N., & Semenov, A. (2025). Smart learning in the 21st century: Advancing constructionism across three digital epochs. Education Sciences, 15(1), 45. https://doi.org/10.3390/educsci15010045
7. Krajcik, J. S., & Shin, N. (2014). Project-based learning. In R. K. Sawyer (Ed.), The Cambridge handbook of the learning sciences (2nd ed., pp. 275–297). Cambridge University Press.
8. Lodi, M., & Martini, S. (2021). Computational thinking, between Papert and Wing. Science & Education, 30(4), 883–908. https://doi.org/10.1007/s11191-021-00202-5
9. Markula, A., & Aksela, M. (2022). The key characteristics of project-based learning: How teachers implement projects in K-12 science education. Disciplinary and Interdisciplinary Science Education Research, 4(1), 2. https://doi.org/10.1186/s43031-021-00042-x
10. Maspul, K. A. (2024). Enhancing STEM education through project-based learning: A strategy to engaging secondary school students. Journal of Education and Learning, 18(2), 312–320.
11. Morales-Morgado, E. M., Ruiz-Torres, S., Rodero-Cilleros, S., Morales-Romo, B., & Campos-Ortuño, R. A. (2023). Metodologías activas en educación superior, mediadas por tecnologías en diversas disciplinas. Aula: Revista de Pedagogía de la Universidad de Salamanca, 29, 67–85. https://doi.org/10.14201/aula202329
12. Papert, S. (1980). Mindstorms: Children, computers, and powerful ideas. Basic Books.

13. Papert, S., & Harel, I. (1991). Situating constructionism. In I. Harel & S. Papert (Eds.), Constructionism (pp. 1–11). Ablex Publishing Corporation.
14. Rehman, A., Bhuttah, T. M., & You, X. (2023). Mapping the global research on project-based learning: A bibliometric and network analysis (2014–2024). Frontiers in Education, 10, 1522694. https://doi.org/10.3389/feduc.2025.1522694
15. Rubin DM, Keene PAC, Richards XL and George A (2026) Can computational modeling in medical education support a constructionist educational framework? Insights from the seminal literature in Papertian constructionism and system dynamics. Front. Educ. 11:1743544. doi: 10.3389/feduc.2026.1743544
16. Silva, M., Correa, R., & Mc-Guire, P. (2024). Metodologías activas con inteligencia artificial y su relación con la enseñanza de la matemática en la educación superior en Chile: Estado del arte. Universidad Nacional de La Plata. https://sedici.unlp.edu.ar/handle/10915/168188
17. Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. Harvard University Press.
18. Benemérita Universidad Autónoma de Puebla. (2009). Modelo Universitario Minerva. BUAP. https://www.minerva.buap.mx
19. Secretaría de Educación Pública. (2023). Marco curricular y Plan de estudios 2022 de la Educación Básica Mexicana: Nueva Escuela Mexicana. SEP.
20. Wokwi. (2024). Wokwi — Simulador de Arduino y ESP32 en línea. https://wokwi.com