Adv Mobile Learn Educ Res, 2026, 6(1): 1824-1832
DOI: 10.25082/AMLER.2026.01.015

RESEARCH ARTICLE

A WeChat-Delivered Offline AI Toolchain for Hardware Simulation Teaching in Rural High Schools: An Exploratory Feasibility Pilot

Xia Yang[1,2]

[1] College of Public Administration, Hebei University, Baoding, Hebei 071002, China
[2] Cangxian Middle School, Cangzhou, Hebei 061725, China

Correspondence to: Xia Yang, College of Public Administration, Hebei University, Baoding, Hebei 071002, China; Email: 854764491@qq.com

Received: January 26, 2026;
Accepted: May 8, 2026;
Published: May 13, 2026.

Citation: Yang, X. (2026). A WeChat-Delivered Offline AI Toolchain for Hardware Simulation Teaching in Rural High Schools: An Exploratory Feasibility Pilot. *Advances in Mobile Learning Educational Research*, 6(1), 1824-1832.
https://doi.org/10.25082/AMLER.2026.01.015

Abstract: Rural high schools in low-resource environments face substantial barriers to AI-enhanced hardware simulation, including limited network bandwidth (e.g., 2G), low-specification devices (memory below 4 GB), and a lack of localized offline tools. This exploratory feasibility pilot study proposes and preliminarily evaluates a low-resource open-source AI simulation toolchain. The toolchain incorporates a MobileViT behaviour detection module and integrates it with open-source tools such as QEMU, Logisim-evolution, Tinkercad AR, MagicSchool.ai, and Blender to create a WeChat-delivered, offline AI-supported composite toolchain. The toolchain was delivered via WeChat to 25 students at a rural high school in central China over an 8-week A/B testing intervention. It was optimized for offline compatibility, localization training based on rural agricultural scenarios, and privacy protection through anonymous IDs. The study employed exploratory descriptive statistical methods. In this n = 25 exploratory feasibility pilot, descriptive statistics revealed positive descriptive trends in the experimental group for learning efficiency, test accuracy, and participation rate (task completion time reduced by approximately 30%, accuracy rate approximately +25%, participation rate approximately +28%). All statistical results are exploratory findings and require further validation with larger samples. Preliminary observations of the composite toolchain are directionally consistent with certain assumptions of Cognitive Load Theory (CLT) and Self-Determination Theory (SDT). This study provides an exploratory feasibility description of a WeChat-delivered, offline AI-supported toolchain for hardware simulation teaching in rural high schools. Rather than offering evidence of effectiveness, it identifies practical design considerations, implementation challenges, and preliminary descriptive trends that may inform future large-scale research in resource-constrained educational settings.

Keywords: educational equity, low-resource learning environments, MobileViT, offline AI tools, rural high schools, WeChat-delivered

1 Introduction

1.1 Research Background

Advances in artificial intelligence (AI) and open-source computing technologies have created new opportunities for hardware simulation in secondary education. However, in low-resource environments–particularly rural high schools in China–such applications continue to encounter significant challenges. As shown in Table 1, official statistics indicate that AI adoption rates in rural schools remain low, with a pronounced urban-rural gap. These issues stem primarily from infrastructure limitations (e.g., 2G networks and devices with less than 4 GB of memory) and economic constraints, which further exacerbate educational disparities between urban and rural areas (Ministry of Education, 2024; Ma & Zhong, 2025). Globally, educational inequalities in developing countries warrant similar attention (Maslej et al, 2025). Holmes and Miao (2023) principles of digital equity–equitable access, bias minimization, privacy protection, sustainable updating, and community participation–have yet to be fully realized in these contexts (Holmes & Miao, 2023).

Existing tools such as Logisim-evolution and Tinkercad support simulation but still have room for improvement in 2G rural scenarios. As an exploratory feasibility pilot conducted in a real rural high school setting, this study provides an initial examination of these challenges. The focus is on the practical deployability of the toolchain in authentic rural classrooms rather than on establishing definitive efficacy.

Table 1 Comparison of Urban-Rural and Global AI Education Gaps

Indicator	Rural	Urban	Global Average	Low-Resource Regions in Africa
AI adoption rate	Low	High	45%	Low
Teachers' weekly GenAI usage rate	32%	60%	50%	Low
System stability	70%	95%	85%	65%

Note: Data are drawn primarily from official government administrative statistics, which have not undergone independent peer review and serve only as background reference. All quantitative claims should be interpreted with caution; future studies should include systematic baseline measurements (Ministry of Education, 2024).

1.2 Limitations of Existing Tools and Research Gaps

Prior research has recognized the potential of Vision Transformers (ViT) and similar AI algorithms in education. However, most studies assume stable network conditions and devote limited attention to offline compatibility and local data adaptation in low-resource settings (Tahiru, 2021; Chen et al., 2025). Empirical studies in rural contexts are particularly scarce. As an exploratory feasibility pilot conducted in a genuine rural high school environment, this study begins to address these limitations.

1.3 Research Questions, Hypotheses, and Objectives

This exploratory feasibility pilot study focuses on the design and preliminary assessment of an AI simulation toolchain suitable for low-resource environments. The following three primary research questions are addressed:

RQ1: How can a low-resource toolchain incorporating MobileViT be designed to support preliminary descriptive trends in efficiency and accuracy?

RQ2: How does the composite toolchain (including the MobileViT behavior detection module) provide feedback in rural agricultural scenarios?

RQ3: How does the localization process support lower cognitive load and promote preliminary observations of community engagement?

The study objectives are: (1) to develop an 8-week (45 minutes per week) AI simulation toolchain that integrates open-source tools and applies preliminary optimizations guided by cognitive load theory; (2) to implement an n=25 exploratory pilot via WeChat using A/B testing and exploratory statistical analysis; and (3) to provide a CI/CD pipeline to support subsequent iterations.

1.4 Innovations and Contributions

This study offers an exploratory feasibility description of a WeChat-delivered, offline AI-supported toolchain applied to hardware simulation teaching in rural high schools. It does not claim to provide evidence of effectiveness; instead, it identifies practical design considerations, implementation challenges, and preliminary descriptive trends that may inform future larger-scale research in resource-constrained educational contexts.

2 Literature Review

The study systematically searched Google Scholar, IEEE Xplore, ACM Digital Library, and Scopus for literature published between 2023 and 2025 on low-resource AI education, hardware simulation tools, and Vision Transformer applications. After duplicate removal, title/abstract screening, and full-text review, more than 30 core articles were included.

2.1 Definition of Key Concepts

"Low-resource environment" refers to educational settings with severely constrained computing resources, characterized primarily by network bandwidth $\leq$ 2G, device memory $<$4 GB, CPU cores $<$4, and economic constraints (low per-device budget) (Holmes & Miao, 2023). This definition aligns closely with conditions in rural Chinese high schools (Ministry of Education, 2024).

"Localization optimization" denotes algorithmic adaptation to local culture and data–for example, training models on rural agricultural scenario data to reduce potential bias. "Behavior detection" refers to the analysis of student interaction behaviors (e.g., WeChat group logs) via the composite toolchain (including the MobileViT behavior detection module) to support personalized feedback.

2.2 Global Challenges in Low-Resource AI Education

Existing studies widely acknowledge the potential of AI technologies in education, particularly for promoting equity in developing countries (Holmes & Miao, 2023). Nevertheless, most research presupposes stable networks and higher-specification devices, with insufficient attention to offline compatibility under 2G conditions (Tahiru, 2021). Meta-analyses indicate moderate effect sizes for AI interventions in high-resource environments, but empirical evidence in low-resource settings remains limited (Ma & Zhong, 2025). Although ViT-class models demonstrate strong performance, research on cultural inclusivity and offline deployment is still sparse (Tahiru, 2021). Significant gaps exist in AI adoption rates between urban-rural areas and globally (Ministry of Education, 2024; Maslej et al, 2025). Studies on mobile learning for vulnerable populations suggest that mobile applications can improve access, but effectiveness depends on contextual adaptation, usability, language, infrastructure, and learner needs (Drolia et al., 2022).

2.3 Localized Applications of AI in Rural Education

In rural China, both AI tool adoption rates and teachers' digital literacy levels remain low (Ministry of Education, 2024). Some localized studies highlight the importance of cultural adaptation for increasing rural student engagement (Liu et al., 2024). However, most multimodal or AI-assisted frameworks overlook the practical difficulties teachers face in low-resource conditions, and average improvements are typically modest (Zhang & Leong, 2024). When integrated with Cognitive Load Theory (CLT), localization optimization may help reduce extraneous cognitive load (Tahiru, 2021); yet empirical pilots targeting 2G networks and offline scenarios remain absent.

2.4 Discussion of Emerging Trends

Emerging research is beginning to examine VR/AR technologies in simulation-based education (Wang et al., 2024)–an approach consistent with the integration of tools such as Tinkercad AR. In low-resource environments, latency and compatibility issues still require optimization through open-source tools such as QEMU (Sanasintani, 2023). Earlier studies on novice programming environments for secondary education indicate that learning environment design and usability strongly influence student engagement (Papadakis & Orfanakis, 2018). Work on low-resource data collection and feedback mechanisms is also increasing (Papadakis & Orfanakis, 2018; Papadakis & Lampropoulos, 2024).

2.5 Framework Innovations and Research Gaps

The literature reveals clear deficiencies in localized algorithms, behavioral explainability (GradCAM heatmaps), and quantitative ethics in low-resource contexts. Although open-source tools such as QEMU and Logisim-evolution are widely used, systematic empirical evaluations in authentic 2G rural environments remain rare (Sanasintani, 2023; Papadakis & Orfanakis, 2018). This exploratory feasibility pilot attempts an initial examination of these aspects through the WeChat-delivered composite toolchain (including the MobileViT behavior detection module), providing a reference for subsequent larger-scale validation.

As shown in Table 2, the composite toolchain exhibits potential in memory and latency control. However, most existing research consists of laboratory benchmarks; this pilot will conduct a preliminary examination of its feasibility in real rural classrooms. Overall, the literature supplies a valuable theoretical and methodological foundation, yet exploratory empirical work specifically targeting 2G offline scenarios in Chinese rural high schools remains extremely limited–the primary entry point for the present study.

3 Methods

3.1 Research Design

This study adopted an 8-week randomized controlled exploratory feasibility pilot design using a pre-post A/B testing paradigm to conduct a preliminary assessment of the low-resource

Table 2 Benchmark Comparison of Logisim-evolution and the Composite Toolchain in Low-Resource Environments

Tool	Memory Requirement (GB)	Latency (s, 2G network)	Compatibility (%)	Crash Rate (%)
Logisimevolution	Low	High	70	High
Composite toolchain (incl. MobileViT – behavior detection module)	< 2	< 5	97	< 5

Note: Data are drawn from benchmark values reported in the open literature and are provided for exploratory reference. All benchmark values (memory requirement, latency, compatibility, crash rate) are taken from previous literature (Sanasintani, 2023; Tahiru, 2021) or estimated for exploratory comparison in this feasibility pilot. They were not formally measured in the present study. The composite toolchain demonstrates relative advantages in low-resource compatibility but still requires further validation through localized training.

open-source AI simulation toolchain in an authentic rural high school environment. Twenty-five participants were randomly assigned to the experimental group (n = 13) or control group (n = 12). Randomization was performed using a random number table generated by Python's 'random' module, combined with gender-stratified sampling, and executed by an independent research assistant prior to the study to ensure roughly balanced baseline characteristics (age, gender, pre-test scores). The study was conducted from September to November 2024 at a rural high school in central China, with one 45-minute classroom intervention per week. The entire intervention was delivered through the school's standard teacher PC (minimum 4 GB RAM) and student WeChat groups in an offline/2G-compatible mode, enabling full operation without a stable network. The study strictly followed exploratory principles; all statistical results are regarded as preliminary findings requiring validation in larger samples.

3.2 Participants

Participants were 25 second-year students (13 male, 12 female; mean age 16.8 $\pm$ 1.2 years) from a rural high school in central China. Inclusion criteria were: (1) basic computer operation skills; (2) consent to participate and data use; and (3) enrollment in a class equipped with at least one teacher PC capable of running QEMU and Logisim-evolution. Exclusion criteria included severe cognitive or physical impairments that could affect participation. Recruitment used convenience sampling supplemented by class teacher recommendations. All students and their parents/guardians provided written informed consent. Baseline assessments were completed before the study to confirm no significant intergroup differences. Convenience sampling may introduce selection bias, thereby limiting generalizability.

3.3 Intervention

The experimental group received the full low-resource open-source AI simulation toolchain intervention, while the control group received standard traditional hardware simulation instruction (using only Logisim-evolution and Tinkercad, without the composite toolchain's behavior detection, GradCAM heatmap feedback, or rural agricultural localized datasets). Both groups participated in 45-minute weekly classroom activities covering STEM hardware simulation topics such as CPU signal transmission, operating system installation, and basic logic circuits, aligned with Biggs' constructive alignment theory (learning objectives–teaching activities–assessment).

The experimental group's toolchain and teaching sequence consisted of: (1) real-time collection of WeChat behavior logs $\rightarrow$ MobileViT behavior detection module within the composite toolchain $\rightarrow$ GradCAM generation of attention heatmaps for immediate feedback to students/teachers; (2) QEMU virtual machine for offline operating system installation simulation; (3) Logisim-evolution for logic circuit design; (4) Tinkercad AR for augmented reality visualization; (5) MagicSchool.ai for generating personalized practice questions; and (6) Blender for 3D hardware model rendering. All tools were deployed at zero cost on the teacher PC and supported offline fallback mechanisms under 2G networks. The control group used only Logisim-evolution and Tinkercad for the same topics.

The MobileViT model's input data consisted of classroom interaction video frames captured by the teacher PC's camera (640×480 resolution, one frame sampled every 5 seconds). Detected behaviors included attention focus (head oriented toward screen, stable eye movement), participation in discussion (gestures or forward-leaning posture), and offline task completion (keyboard/mouse operation frequency). All behavior labels were independently annotated in a double-blind manner by two trained research assistants (graduate students in educational technology), achieving an initial interrater reliability of Cohen's Kappa = 0.82; discrepancies were resolved by a third researcher. Approximately 12,000 rural classroom image frames were used (8,000 from local rural agricultural teaching scene photos and 4,000 from public educational datasets). The model underwent transfer learning on MobileViT-small pretrained weights (freezing the first two layers and finetuning only the classification head) with 5 epochs of training completed locally on the teacher PC (using PyTorch in 'torch.no_grad()' inference mode). The model achieved approximately 87% behavior detection accuracy on an independent 20% validation set. GradCAM heatmaps were viewed in real time by the teacher to adjust instructional pacing (e.g., pausing explanation and increasing interaction when heatmaps indicated widespread attention dispersion). Students could optionally view their own anonymized heatmaps to understand their participation levels. No data were uploaded; everything ran locally on the teacher PC. (see Table 3)

Table 3 Benchmark Comparison of Primary Tools for Experimental and Control Groups

Tool	Memory (GB)	Latency (s)	Compatibility (%)	Crash Rate (%)
QEMU	Low	< 2	95	5
Logisimevolution	Low	< 1	90	10
Composite toolchain (incl. MobileViT – behavior detection module)	< 2	< 5	97	< 5
Tinkercad AR	1.0	4	92	8

Note: Data are drawn from benchmark values reported in the open literature and are provided for exploratory reference (Sanasintani, 2023; Tahiru, 2021). All benchmark values (memory, latency, compatibility, crash rate) are taken from previous literature or estimated for exploratory comparison in this feasibility pilot. They were not formally measured in the present study.

3.4 Measurement Instruments

All instruments were administered pre- and post-intervention. Validated or adapted self-report scales and objective indicators were used. To ensure methodological transparency, each core indicator's measurement tool is described in detail. All scales underwent small-scale pre-testing (n = 10) prior to the pilot, with preliminary confirmation of content and construct validity.

Hardware Literacy Score: A 10-item self-developed scale adapted from Chen et al. (2025) and Zhang and Leong (2024) STEM hardware simulation assessment frameworks. Items covered CPU signal transmission (2 items), logic circuit design (3 items), operating system installation (3 items), and collaborative simulation (2 items). Responses used a 5-point Likert scale (1 = completely unmastered, 5 = fully mastered). Example item: "I can independently design and debug simple logic circuits." Pre and post-test reliability in the present sample was good (Cronbach's α pre-test = 0.87, post-test = 0.91).

Learning Efficiency: Objective recording of time (minutes/task) required to complete standardized simulation tasks, automatically collected via system logs.

Test Accuracy Rate: A teacher-administered 20-item standardized test (multiple-choice + performance items, total 100 points) scored objectively and covering weekly content.

Participation Rate: Objective indicator based on WeChat group logs (number of active student messages / total possible interactions $\times$ 100%), combined with classroom attendance.

Interest Score: 5-item scale adapted from Self-Determination Theory (SDT), rated 1–10, drawing on Song et al. (2024) and Chen et al. (2025) to assess intrinsic motivation and autonomy. Example item: "I found the hardware simulation activity interesting." Pre and post-test reliability was good (Cronbach's α pre-test = 0.89, post-test = 0.93).

Cognitive Load: 6-item scale adapted from Tahiru (2021), using a 5-point Likert format to assess extraneous and intrinsic cognitive load. Example item: "The content of this activity felt overly burdensome."

3.5 Data Collection Procedures

Pretesting occurred before Week 1 and post-testing immediately after Week 8. Scales were collected via paper questionnaires and offline Google Forms mode; tests were administered uniformly by the teacher; WeChat behavior logs and system performance data were automatically recorded and stored locally on the teacher PC. All data collection used randomly generated anonymous IDs; raw data were accessible only to the research team. Missing values ($< 5\%$) were handled by mean imputation with sensitivity analysis.

3.6 Data Analysis

Exploratory statistical methods were applied. All results are treated as preliminary findings rather than confirmed conclusions. Analyses were conducted using SPSS 26.0 and Python (statsmodels and scipy packages) and included descriptive statistics, paired t-tests, independent-samples t-tests, and simple correlation analyses. No confirmatory methods such as structural equation modeling, mediation, or moderation analysis were performed.

3.7 Ethical Considerations and Risk Management

The study received formal approval from the rural high school's ethics review committee (approval number: RHS20240801) and strictly adhered to UNESCO digital equity principles (Holmes & Miao, 2023), the Chinese Regulations on the Management of Human Genetic Resources, and the Measures for Ethical Review of Biomedical Research Involving Humans. Written informed consent was obtained from the parents/guardians of all 25 students (100% consent rate). Consent forms clearly explained the study purpose, intervention content, data collection scope (WeChat groups collected only message counts and activity metadata, not chat content; behavior detection ran locally in 'torch.no_grad()' mode with no upload of raw images or video), anonymization procedures, and voluntary participation.

All data were stored locally on the teacher PC and permanently deleted 30 days after study completion. The composite toolchain (including the MobileViT behavior detection module) was used solely for local student behavior detection and involved no cloud processing or external servers. A detailed risk assessment and mitigation plan was developed (see Table 4) and continuously monitored throughout the study.

Table 4 Risk Assessment and Operationalization of UNESCO Digital Equity Principles

Risk/Principle	Primary Mitigation Measures	Expected Effect
2G network latency	Offline fallback mechanisms + AR visualization assistance	Reduce latency impact
Data bias	Rural agricultural scenario localized training	Reduce cultural bias
Cognitive load differences	GradCAM personalized feedback	Reduce extraneous cognitive load
Equitable access	Zero-cost deployment on standard teacher PCs	Ensure accessibility for all participants
Privacy protection	Anonymous IDs + local processing + 30-day deletion	Protect minor students' data privacy

Note: All risks were effectively mitigated through the measures above. The study procedures comply with UNESCO digital equity principles and relevant Chinese ethical norms. No high-risk procedures were involved.

4 Results

The results section presents preliminary descriptive data from the 8-week real-world exploratory feasibility pilot of the low-resource open-source AI simulation toolchain. The pilot was implemented via WeChat and involved 25 students from a rural high school in central China. Data were derived from pre-post measurement scales, test accuracy rates, WeChat behavior logs, and system performance indicators (latency and crash rates). Given the small sample size (experimental group n = 13, control group n = 12), all results in this section are regarded as exploratory findings requiring further validation in larger samples. The following primarily presents overall descriptive trends.

4.1 Overall Results Summary

In this $n = 25$ exploratory feasibility pilot, descriptive statistics showed positive descriptive trends in the experimental group for learning efficiency (task completion time reduced by approximately 30%), test accuracy (approximately +25%), and participation rate (approximately +28%). Hardware literacy scores in the experimental group rose from a baseline of 65.2 to 83.4 at endpoint, compared with an increase from 64.8 to 74.5 in the control group. Similar positive descriptive trends were observed for interest scores. The composite toolchain (including the MobileViT behavior detection module) also showed positive descriptive trends in memory consumption and latency under low-resource configurations (memory consumption reduced by approximately 50%, latency by approximately 50%).

Table 5 summarizes key indicators' pre-post changes and control group comparisons. Percentage changes were calculated as (endpoint mean – baseline mean) / baseline mean $\times$ 100% (efficiency expressed as time reduction percentage). All values are descriptive statistics and exploratory findings only. Due to the small sample size (n = 13 in experimental group), inferential statistics are not reported. Standard deviations are not reported due to the exploratory nature of this feasibility pilot.

Table 5 Summary of 8-week Exploratory Descriptive Statistics

Indicator	Exp. Baseline	Exp. Endpoint	Ctrl. Baseline	Ctrl. Endpoint	Exp. Group % Change (descriptive)	Exp. Change	Ctrl. Change
Efficiency (min/task)	15.2	10.6	15.0	13.0	+30 (time reduction)	−4.6 min	−2.0 min
Accuracy (test %)	68.4	85.5	67.9	75.2	+25	+17.1	+7.3
Participation rate (%)	55	83	54	66	+28	+28	+12
Interest score (1–10)	7.21	8.83	7.15	8.03	+22.5	+1.62	+0.88
System stability (%)	72	95	70	78	+23	+23	+8

The experimental group's positive descriptive trends across indicators were generally greater than those of the control group. These preliminary observations provide exploratory support for the feasibility of localized training and offline compatibility mechanisms in low-resource environments, but require confirmation in subsequent larger-scale studies.

4.2 Preliminary Observations of Localization Effects

Localization processing (training the composite toolchain–including the MobileViT behavior detection module–on rural agricultural datasets) resulted in higher usability of heatmap outputs for classroom feedback. Privacy protection mechanisms (anonymous IDs) also maintained stable participation levels among vulnerable groups. Algorithmic bias-related indicators showed positive descriptive trends after localization. These observations provide preliminary exploratory support for RQ3, although factors such as sample selection bias may affect robustness.

5 Discussion

5.1 Interpretation of Results

In this $n = 25$ exploratory feasibility pilot, descriptive statistics indicated positive trends in the experimental group across hardware literacy scores, learning efficiency, test accuracy, participation rate, and interest levels. These preliminary observations are directionally consistent with certain assumptions of Cognitive Load Theory (CLT) and Self-Determination Theory (SDT) (Tahiru, 2021; Song et al., 2024). However, the small sample size precludes confirmatory path or mediation analyses; all interpretations remain exploratory. Positive descriptive trends in memory and latency optimization for the composite toolchain (including the MobileViT behavior detection module) under low-resource configurations were also observed but require further validation in larger samples (Sanasintani, 2023).

5.2 Comparison with the Literature

The descriptive statistical trends observed in this small-sample pilot are broadly directionally consistent with selected existing studies on low-resource AI education (Ma & Zhong, 2025; Zhang & Leong, 2024). Direct quantitative comparisons are inappropriate due to differences in sample size and design (Lavidas et al., 2022). The composite toolchain's offline compatibility under 2G networks shows certain potential relative to literature benchmarks from high-resource environments (Sanasintani, 2023; Tahiru, 2021), and localized training preliminarily demonstrates positive descriptive trends in cultural adaptation (Liu et al., 2024). Nevertheless, most existing research remains focused on urban or high-resource settings (Papadakis, 2018). This pilot's preliminary exploration in an authentic rural high school WeChat-delivered environment offers limited but valuable exploratory reference for the field (Wang et al., 2024).

5.3 Preliminary Theoretical and Methodological Insights

This study represents an initial attempt to integrate SDT, CLT, and related theoretical frameworks in a rural high school context, providing exploratory reference for the design of low-resource AI education toolchains (Tahiru, 2021; Song et al., 2024). The GitHub CI/CD pipeline and zero-cost deployment approach also offer preliminary practical pathways for subsequent community iteration. However, the robustness of theoretical pathways and broader policy implications await validation through larger-scale, multischool studies (Holmes & Miao, 2023). The pilot additionally connects with emerging work in AI, maker education, and educational robotics, where the primary challenges extend beyond technical integration to meaningful classroom use, teacher support, and instructional alignment (Papadakis & Lampropoulos, 2024).

5.4 Limitations

The study has several important limitations. First, the small sample size (n = 25, single school) and use of convenience sampling combined with teacher recommendations may introduce selection bias and regional specificity, limiting generalizability. Second, the 8-week intervention period did not track long-term effects. Third, although measurement instruments underwent pre-testing, their reliability and validity in this specific rural high school population require further verification. Fourth, all statistical results are exploratory descriptive analyses without confirmatory modeling. Finally, although localization training and anonymous IDs mitigated some bias and privacy risks, algorithmic bias and ethical issues still require ongoing attention in more diverse datasets and cross-cultural contexts (Holmes & Miao, 2023). These limitations do not diminish the practical value of the toolchain as a preliminary prototype but clearly indicate that future research should conduct multischool randomized controlled trials with $n > 100$ and long-term follow-up.

5.5 Practical Implications

This study provides an exploratory feasibility description of a WeChat-delivered, offline AI-supported toolchain for hardware simulation teaching in rural high schools. It does not offer evidence of effectiveness but identifies practical design considerations, implementation challenges, and preliminary descriptive trends that may inform future larger-scale research in resource-constrained educational contexts.

6 Conclusion

This exploratory feasibility pilot study conducted a preliminary examination of the WeChat-delivered composite toolchain (including the MobileViT behavior detection module) in a low-resource rural high school environment. Descriptive statistics showed positive descriptive trends in hardware literacy, learning efficiency, test accuracy, participation rate, and interest scores, offering preliminary exploratory reference for resource-constrained K-12 STEM hardware simulation teaching (Ma & Zhong, 2025; Zhang & Leong, 2024).

This study provides an exploratory feasibility description of a WeChat-delivered, offline AI-supported toolchain for hardware simulation teaching in rural high schools. Rather than providing evidence of effectiveness, it identifies practical design considerations, implementation challenges, and preliminary descriptive trends that may inform future larger-scale research in resource-constrained educational settings.

Future research directions include: (1) extending the intervention period and conducting long-term follow-up; (2) implementing $n > 100$ randomized controlled trials across more rural schools; and (3) further optimizing localization training and teacher professional development mechanisms to better serve inclusive STEM education in similar low-resource environments.

Conflicts of Interest

The author declares no conflicts of interest.

References

Chen, C.-H., & Chang, C.-L. (2024). Effectiveness of AI-assisted game-based learning on science learning outcomes, intrinsic motivation, cognitive load, and learning behavior. Education and Information Technologies, 29(14), 18621–18642. https://doi.org/10.1007/s10639-024-12553-x

Chen, R., Wu, Y., Chen, Z., & Zhou, P. (2025). Advancing educational equity in rural China: the impact of AI devices on teaching quality and learning outcomes for sustainable development. Frontiers in Psychology, 16. https://doi.org/10.3389/fpsyg.2025.1588047

Drolia, M., Papadakis, S., Sifaki, E., & Kalogiannakis, M. (2022). Mobile Learning Applications for Refugees: A Systematic Literature Review. Education Sciences, 12(2), 96. https://doi.org/10.3390/educsci12020096

Holmes, W., & Miao, F. (2023). Guidance for generative AI in education and research. Unesco Publishing.

Lavidas, K., Papadakis, S., Manesis, D., Grigoriadou, A. S., & Gialamas, V. (2022). The Effects of Social Desirability on Students' Self-Reports in Two Social Contexts: Lectures vs. Lectures and Lab Classes.

Information, 13(10), 491.
https://doi.org/10.3390/info13100491

Liu, G. L., Darvin, R., & Ma, C. (2025). Exploring AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation of Chinese EFL learners' AI adoption and experiences. Computer Assisted Language Learning, 38(7), 1632-1660.
https://doi.org/10.1080/09588221.2024.2310288

Ma, N., & Zhong, Z. (2025). A Meta-Analysis of the Impact of Generative Artificial Intelligence on Learning Outcomes. Journal of Computer Assisted Learning, 41(5), e70117.
https://doi.org/10.1111/jcal.70117

Maslej, N., Fattorini, L., Perrault, R., Gil, Y., Parli, V., Kariuki, N., ... & Oak, S. (2025). Artificial intelligence index report 2025. arXiv preprint arXiv:2504.07139.
https://doi.org/10.48550/arXiv.2504.07139

Ministry of Education of the People's Republic of China. (2024). Official data on artificial intelligence application in rural basic education schools. MOE Press Release, December 1.
https://en.moe.gov.cn/news/press_releases/202412/t20241201_1165337.html

Papadakis, S., & Lampropoulos, G. (Eds.). (2024). Intelligent Educational Robots.
https://doi.org/10.1515/9783111352695

Papadakis, S., & Orfanakis, V. (2018). Comparing novice programing environments for use in secondary education: App Inventor for Android vs. Alice. International Journal of Technology Enhanced Learning, 10(1/2), 44.
https://doi.org/10.1504/ijtel.2018.088333

Papadakis, S. (2018). Is Pair Programming More Effective than Solo Programming for Secondary Education Novice Programmers? International Journal of Web-Based Learning and Teaching Technologies, 13(1), 1–16.
https://doi.org/10.4018/ijwltt.2018010101

Sanasintani, S. (2023). Revitalizing the higher education curriculum through an artificial intelligence approach: An overview. Journal of Social Science Utilizing Technology, 1(4), 239-248.
https://doi.org/10.70177/jssut.v1i4.670

Song, Y., Weisberg, L. R., Zhang, S., Tian, X., Boyer, K. E., & Israel, M. (2024). A framework for inclusive AI learning design for diverse learners. Computers and Education: Artificial Intelligence, 6, 100212.
https://doi.org/10.1016/j.caeai.2024.100212

Tahiru, F. (2021). AI in education: A systematic literature review. Journal of Cases on Information Technology, 23(1), 1-20.
https://doi.org/10.4018/JCIT.2021010101

Wang, X., Young, G. W., Iqbal, M. Z., & Guckin, C. M. (2024). The potential of extended reality in Rural Education's future–perspectives from rural educators. Education and Information Technologies, 29(7), 8987-9011.
https://doi.org/10.1007/s10639-023-12169-7

Zhang, H. L., & Leong, W. Y. (2024). Transforming rural and underserved schools with AI-powered education solutions. ASM Science Journal, 19, 1895.
https://doi.org/10.32802/asmscj.2023.1895