TechMecha RoboSphere | 2026, Vol. 2, No.1, 1 – 12 | ISSN Online: 3155-6019
DOI: https://doi.org/10.62718/vmca.tech-robo.2.1.SC-0126-018

OPEN ACCESS

Development and Pilot Evaluation of a Portable Ethical Hacking Toolkit for Cybersecurity Education

Article History:

Initial submission:	08 October 2026
First decision:	15 October 2026
Revision received:	10 November 2026
Accepted for publication:	18 November 2026
Online release:	21 November 2026

John Paul F. Avila[1], ORCID No. 0009-0000-6915-8058
Nelson Rodelas[2], ORCID No. 0000-0002-6138-4488

[1]Instructor I, Computer Engineering Department, Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines
[2]Faculty, Computer Engineering Department, Polytechnic University of the Philippines, Sta. Mesa, Manila, Philippines

Abstract

Cybersecurity courses in many Philippine state universities rely heavily on lectures and software-based exercises. In this study, Computer Engineering students at the Polytechnic University of the Philippines (PUP) rated access to hardware-based tools and laboratory resources lowest among the instructional conditions surveyed. We developed and pilot-tested a portable ethical hacking toolkit built on a master–satellite network of three ESP32 microcontrollers with NRF24L01 radio modules, intended for supervised classroom use. A developmental-descriptive design with an iterative design model (needs analysis, design, prototyping and development, testing and pilot evaluation, refinement and iteration) guided the work, and the toolkit was piloted over three sessions with one Computer Engineering section. Five closed-ended, six-point Likert instruments gathered perceptions from students before (n = 42) and after (n = 38) the pilot, teaching faculty (n = 10), IT and network-security professionals (n = 13), and faculty and administrators (n = 10). Post-pilot and expert composite means ranged from 5.60 to 5.91 ("Strongly Agree"); students' pre-pilot ratings of hardware access and laboratory resources were 4.19 ("Slightly Agree"). These are perception data from a single section, without a comparison group, an objective learning measure or inferential testing. They indicate acceptability, perceived curriculum alignment and perceived feasibility, not demonstrated learning gains. The findings justify a larger evaluation that includes objective competency assessment, multiple sections and institutions, and costed implementation.

Keywords: ethical hacking, cybersecurity education, ESP32, experiential learning, iterative design, pilot evaluation

Copyright @ 2026. The Author/s. Published by VMC Analytiks Multidisciplinary Journal News Publishing Services. Development And Pilot Evaluation of a Portable Ethical Hacking Toolkit for Cybersecurity Education © 2026 by John Paul F. Avila and Nelson Rodelas is an open access article licensed under Creative Commons Attribution (CC BY 4.0). This permits the copying, redistribution, remixing, transforming, and building upon the material in any medium or format for any purpose, even commercially, provided that appropriate credit is given to the copyright owner/s through proper and standard citation.

INTRODUCTION

Cybersecurity competence is demonstrated by doing configuring networks, observing how attacks unfold, and reasoning about defenses. Ethical hacking, the authorized probing of systems to find weaknesses, is a common vehicle for teaching these skills. Yet cybersecurity education still leans heavily on lectures, workshops, seminars and conventional laboratories (Katsantonis et al., 2023). A recent Philippine study that combined job-listing analysis with surveys of students, teachers and professionals described a gap between academic preparation and industry expectations and thus recommended more practical experiences in programs (Miranda et al., 2025).

State universities and colleges (SUCs) in the Philippines work under particular constraints. A multiple-case study of selected SUCs in the National Capital Region identified user education, cloud security, information-security strategy and unsecured personal devices as recurring cybersecurity challenges, and called for a program framework to guide planning, implementation and assessment (De Ramos & Esponilla, 2022). Studies of technology adoption in SUCs likewise report slow internet connectivity and limited awareness as barriers (Alimboyong & Bucjan, 2021).

At the Polytechnic University of the Philippines (PUP), cybersecurity and network-security courses have been introduced to the Computer Engineering program only within roughly the

last five years, under the program's national policies, standards and guidelines (Commission on Higher Education [CHED], 2017). The authors' working premise was that instruction in these courses is mostly theoretical or simulation-based, and that hardware-based exercises are limited by budget and laboratory resources. This paper treats that premise as something to be tested, through a needs analysis and a pre-implementation student survey, rather than as an established fact.

Existing solutions offer useful precedents. Cyber ranges (Katsantonis et al., 2023) and scalable virtual learning environments (Vykopal et al., 2021) provide authentic and safe practice, although the latter authors note that preparing and delivering hands-on classes usually does not scale, and their environments emulate rather than expose physical devices. Low-cost embedded platforms offer another route: ESP32-based projects have been used to build confidence and data literacy in engineering students (Zambrano-Mieles et al., 2026), and a hands-on cyber-physical security lab built on accessible hardware was well received by students, most of whom had never completed such a lab (Chi et al., 2022). In the literature reviewed for this study, we found few reports of a portable, low-cost toolkit for supervised ethical-hacking exercises designed for resource-constrained SUCs and accompanied by evidence on how it was built, how users perceived it and whether its adoption is realistic.

This paper therefore reports the development and pilot evaluation of a Portable Ethical Hacking Toolkit (PEHT) built on a master–satellite network of ESP32 microcontrollers. It addresses six research questions (RQs):

1. What limitations in hands-on and hardware-supported learning do Computer Engineering students perceive in current cybersecurity instruction?

2. How do faculty rate the curriculum alignment of the toolkit's design?

3. How was the toolkit developed through an iterative design model, and how do experts rate that process?

4. How do students and faculty perceive the toolkit's effect on practical understanding and engagement?

5. How do they perceive its usability, scalability and sustainability?

6. Do faculty and administrators judge the proposed implementation plan feasible?

The study does not claim causal learning effects or generalizability; it documents a pilot and the perceptions it produced.

LITERATURE REVIEW

Hands-on learning in cybersecurity. Experiential approaches are attractive because they let learners practice with realistic scenarios. Chi et al. (2022) described a low-cost, hands-on cyber-physical security lab and tested it with 21 university students using a pretest, posttest and satisfaction survey; most students had not previously done a hands-on lab in the field, and they rated enjoyment highly.

Zambrano-Mieles et al. (2026) reported a sixteen-week project-based intervention with 91 first-year engineering students using ESP32 microcontrollers and web dashboards, in which perceived learning confidence rose from 3.9 to 4.6 on a five-point scale between the challenge and reflection phases. Neither study concerns wireless ethical hacking, and both rely on small samples or self-reported measures, but together they suggest that inexpensive hardware can support structured hands-on teaching in engineering programs. Criollo-C et al. (2024) add an evaluation lens: with 60 engineering students they combined the IBM Computer System Usability Questionnaire (CSUQ) and the NASA Task Load Index (NASA-TLX) and found that a gamified cybersecurity application supported intention to use while imposing low mental workload.

Cyber ranges and scalable environments. Katsantonis et al. (2023) proposed a Cyber Range Design Framework, comprising an architecture and a life cycle built on a conceptual framework for e-learning and training, to guide the design of cyber ranges for education. Vykopal et al. (2021) described interactive learning environments delivered on-site or remotely to school, university and professional learners, motivated by the observation that hands-on classes are costly to prepare and deliver at scale. These resources show what well-designed environments contain: defined scenarios, learning objectives and reusable components. They emulate networks and devices in software, however, and may be difficult for programs with limited infrastructure to adopt directly.

Competency alignment. The NICE Workforce Framework describes cybersecurity work in terms of tasks, knowledge and skills (Petersen et al., 2020), and Karagiannis et al. (2024) designed a cyber range in which scenarios are mapped to NICE work roles. Such mapping lets instructors convert hands-on exercises into observable learning outcomes. In the Philippine setting, a toolkit must also fit the policies, standards and guidelines of the Computer Engineering program (CHED, 2017).

Theoretical and design foundations. Kolb's (1984) experiential learning theory treats learning as a cycle of concrete experience, reflective observation, abstract conceptualization and active experimentation. The toolkit's session structure of demonstration, practice and debrief was designed to traverse this cycle. Development followed iterative design logic. Peffers et al. (2007) describe design science research as problem identification, definition of objectives, design and development, demonstration, evaluation and communication, with iteration when evaluation reveals shortcomings. McKenney and Reeves (2019) frame educational design research as recurring phases of analysis and exploration, design and construction, evaluation and reflection, and implementation and spread, carried out in real instructional settings. The CDIO approach organizes engineering education around conceiving, designing, implementing and operating real artifacts (Crawley et al., 2007). The five stages used in this study map onto these frameworks: needs analysis (problem identification; conceive), design and prototyping (design and development; design and implement), testing and pilot evaluation (demonstration and evaluation; operate), and refinement and iteration (iteration; evaluation and reflection).

Implementation and sustainability. Innovations spread when they offer relative advantage, compatibility, low complexity, trialability and observability (Rogers, 2003). Perceived usefulness and ease of use shape acceptance (Davis, 1989), and facilitating conditions, the belief that organizational and technical support exists, are a distinct determinant of use (Venkatesh et al., 2003). At the organizational level, Fixsen et al. (2005) describe implementation as passing through exploration, installation, initial implementation and full implementation, and caution that the process takes longer than organizations expect. Coburn (2003) argues that scale should be judged by depth, sustainability, spread and shift in ownership rather than by the number of adopting sites. In Philippine SUCs, connectivity and awareness have been reported as adoption barriers (Alimboyong & Bucjan, 2021), and acquisitions are governed by public-procurement rules (Republic Act No. 9184, 2003), which makes procurement lead time a practical planning concern.

Synthesis and conceptual framework. The literature supports three design commitments: practice should be hands-on and scenario-based; activities should be aligned with recognized competency frameworks and local program outcomes; and adoption should be planned as a staged process with institutional ownership. It leaves open whether a portable, low-cost, ESP32-based toolkit can meet these commitments in a resource-constrained Philippine SUC. The study's conceptual framework follows an input–process–output

(IPO) logic. Inputs are the identified instructional gap, the selected hardware platform and the ethical-hacking modules chosen for course outcomes. The process is iterative development and classroom piloting. Outputs are a functional toolkit, guided exercises and evidence on perceived alignment, engagement, usability and feasibility, which feed the next development cycle.

METHODS

Research design. The study used a developmental-descriptive design. The developmental component was the design, prototyping and piloting of the toolkit through the iterative design model described above. The descriptive component was a quantitative, perception-based evaluation using closed-ended surveys. No comparison group was used, and no objective measure of student performance was collected.

Setting. The study was conducted in the Computer Engineering program of PUP. The pilot took place in a Networks and Security course section during the second semester of School Year 2025–2026.

Table 1

Development stages, activities, outputs and decision criteria

Stage	Period	Activities	Outputs	Decision criteria
1. Needs analysis	Dec 2025	Literature review; assessment of the instructional gap at PUP; review of program requirements (CHED, 2017)	Hardware-based exposure identified as the shortfall; ESP32 and NRF24L01 modules selected	Low cost, portability, wireless capability, open-source compatibility; compatibility with mandated outcomes
2. Design	Dec 2025–Jan 2026	Designed a master–satellite architecture of three ESP32 units; organized content into five instructional groupings; defined safeguards	System architecture; module map; restriction of the radio subsystem to pre-authorized, isolated test devices	Multi-channel 2.4 GHz coverage without overloading the master; each grouping mapped to an ethical-hacking phase; design must not allow unauthorized scanning
3. Prototyping and development	Jan–Feb 2026	Hardware assembly; firmware for master and satellite units; bench testing, revision and re-testing; refinement of the onboard interface	Working prototype with three-button input, TFT display and microSD logging	Functional limitations found in bench tests, especially master–satellite coordination, resolved; interface usable without a laptop
4. Testing and pilot evaluation	Feb–Jul 2026	Pre-implementation survey; three classroom sessions; post-implementation survey	Student pre- and post-pilot data; faculty observations	Sessions run only within authorized, supervised, isolated settings
5. Refinement and iteration	Mar–Jul 2026	Faculty survey; expert validation of Stages 1–4; feasibility survey of the implementation plan; review of chairperson comments	Validation ratings; four-phase implementation plan; inputs for the next cycle	Ratings and comments treated as inputs to the next iteration rather than as closing judgments

Note. *Stages overlapped because development was iterative: feedback gathered during testing, the pilot and expert review was incorporated into earlier design and prototyping work. Periods show when each stage was active, not strict boundaries.*

Development stages. Development ran from December 2025 to July 2026 in five interlinked stages. Table 1 lists the activities, outputs and decision criteria of each. The needs analysis combined a review of the literature, a direct assessment of the instructional gap in the PUP program, and a review of the program's policies, standards and guidelines (CHED, 2017). The student pre-implementation survey (reported under RQ1) was added as a direct measure of the gap before any student used the toolkit. Recommendations from faculty, experts and administrators received during later stages were fed back into the design, so the stages overlap and are listed in the order each took the lead.

System overview. The toolkit is powered through USB (battery pack or charger) and operated through a three-button input (Preview, Select, Next) with a 3.5-inch SPI TFT display and a microSD module that logs session data for debriefs. Three ESP32 units form a master–satellite network: the master handles navigation and display and aggregates data from two satellites, each interfaced to three NRF24L01 2.4 GHz transceivers, which lets the toolkit monitor several channels at once. Its Wi-Fi and Bluetooth functions extend conventions from open-source ESP32 wireless-testing firmware.

The hardware cost is approximately PHP 3,500 to 4,000 per complete unit. Content is organized into five instructional groupings, each tied to an ethical-hacking phase: Device Overview; Wi-Fi-Based Attacks (reconnaissance, beacon spamming, probe flooding, deauthentication); Credential Harvesting (rogue access point and captive-portal pages); Bluetooth-Based Attacks (BLE advertisement spamming and HID keystroke-injection demonstration); and 2.4 GHz Attacks (carrier transmission, passive scanning and a channel-activity analyzer). The pilot closed with a combination exercise in which students launched two attacks together. Each grouping was followed by discussion of detection and mitigation. The radio subsystem was limited to a fixed set of pre-authorized, isolated test devices.

Participants and selection. Four groups took part (Table 2). Students were all 42 enrolled members of one intact section; 38 were present for the post-implementation survey. The other groups were purposively selected for relevant exposure or expertise.

Table 2
Respondent groups and selection basis

Group	n	Role in the study	Selection basis
Students (pre / post)	42 / 38	Completed baseline and post-pilot surveys	All students enrolled in the pilot section, 2nd semester, SY 2025–2026 (single intact class)
Teaching faculty	10	Observed or co-facilitated the pilot; completed the Faculty Post-Implementation Survey	Tenured professors teaching networks and cybersecurity courses who attended the pilot sessions.
IT and network-security professionals	13	Reviewed development documentation; completed the Development Process Validation Checklist	Cybersecurity and Networking industry practitioners with experiences ranging from 3 years to 15 or more years.
Faculty and administrators	10	Rated the proposed implementation plan	Faculty, program leaders or administrators knowledgeable about implementing cybersecurity courses and department resources at PUP including Program Chairperson, former laboratory heads, and staff.

Note. Group sizes follow the instruments administered.

Instruments. Five closed-ended instruments used a six-point Likert scale from Strongly Disagree to Strongly Agree: (1) Student Pre-Implementation Survey (existing limitations; baseline readiness); (2) Student Post-Implementation Survey (engagement and practical understanding; usability and scalability); (3) Faculty Post-Implementation Survey (curriculum alignment; observed engagement; usability, scalability and sustainability); (4) Toolkit Development Process Validation Checklist (15 indicators across the five stages); and (5) Implementation Plan Feasibility Survey (12 indicators in three domains). Academic experts reviewed all instruments for content validity before administration (ten faculty members with networking backgrounds, current or former administrators at the university). Reliability coefficients are not reported in this paper.

Procedure. Session 1 covered device orientation and Wi-Fi-based exercises; Session 2, credential harvesting and Bluetooth-based exercises; Session 3, 2.4 GHz exercises and the combination exercise. Each followed a demonstration–practice–debrief structure aligned with Kolb's (1984) cycle. Students completed the pre-implementation survey before any exposure and the post-implementation survey at the end of Session 3. Faculty completed their survey after observing or co-facilitating the sessions. The expert panel and the faculty–administrator group completed their instruments separately from the classroom sessions, after reviewing the development documentation and the implementation plan, respectively. The plan has four phases: pilot implementation (completed); installation and departmental onboarding, including training of additional faculty trainers, a maintenance budget and procurement lead time; curricular integration under the program's policies, standards and guidelines (CHED, 2017); and full implementation with periodic monitoring.

Data analysis. Responses were summarized as weighted means and standard deviations per item and per construct using spreadsheet software and interpreted with the scale in Table 3. Because pre- and post-pilot items differ and responses were anonymous, no inferential test of change was computed.

Table 3
Interpretation of weighted means on the six-point scale

Mean range	Interpretation
5.16–6.00	Strongly Agree
4.33–5.15	Agree
3.50–4.32	Slightly Agree
2.67–3.49	Slightly Disagree
1.84–2.66	Disagree
1.00–1.83	Strongly Disagree

Ethical considerations. The study received Ethics Review Clearance from the PUP Graduate School Research and Extension Office (Clearance No. 2026-377). Participation was voluntary and anonymous, and no personally identifiable information was collected. Students were briefed on the study's purpose and on the ethical boundaries of toolkit use before the first session. All exercises used pre-authorized, isolated test devices under direct faculty supervision, with no interaction with production networks, and 2.4 GHz activity was confined to the study's radio-frequency authorization.

RESULTS

Table 4 summarizes weighted means by research question; item-level results follow. All data are self-reported perceptions, so the results describe what respondents believed, not independently measured performance.

Table 4

Summary of weighted means by research question

RQ	Construct	Respondents	n	M	Interpretation
1	Existing limitations (adequacy items; gap item)	Students (pre)	42	4.19–5.10; 5.02	Slightly Agree–Agree
2	Curriculum-aligned design	Faculty	10	5.85	Strongly Agree
3	Development process	IT and security professionals	13	5.91	Strongly Agree
4	Practical understanding and engagement (post-pilot)	Students; faculty	38; 10	5.60; 5.72	Strongly Agree
5	Usability; scalability and sustainability	Students; faculty	38; 10	5.83; 5.81	Strongly Agree
6	Implementation-plan feasibility	Faculty and administrators	10	5.78	Strongly Agree

Note. M = weighted mean.

RQ1. Existing limitations. Students agreed that existing courses provide sufficient hands-on activity (M = 5.10, SD = 0.91) and bridge theory and practice (M = 4.55, SD = 1.31). They gave the lowest ratings to opportunities to use hardware-based tools (M = 4.19, SD = 1.70) and to the adequacy of laboratory resources (M = 4.19, SD = 1.55), both "Slightly Agree" (Table 5). They also agreed that a significant gap exists between classroom knowledge and industry skills (M = 5.02, SD = 1.42). The large standard deviations on the hardware items show that views varied widely.

Table 5

Existing limitations in cybersecurity instruction (Student Pre-Implementation Survey, Part I, n = 42)

Indicator	M	SD	Interpretation
1. Courses provide sufficient hands-on, practical activities	5.10	0.91	Agree
2. Adequate opportunities to use hardware-based tools	4.19	1.70	Slightly Agree
3. Laboratory resources meet practical learning needs	4.19	1.55	Slightly Agree
4. Confident applying concepts to real-world scenarios	4.33	1.54	Agree
5. Instruction bridges theory and practice	4.55	1.31	Agree
6. Significant gap between classroom knowledge and industry skills	5.02	1.42	Agree

Note. Item 6 is worded in the opposite direction (agreement indicates a gap), so a single composite across items 1–6 is not interpreted here.

RQ2. Curriculum-aligned design. Faculty rated all six alignment indicators between 5.80 and 5.90 (composite M = 5.85, SD = 0.36; Table 6). The department chairperson, the only faculty respondent who gave an open-ended comment, listed topics for future integration: network security, ethical hacking and penetration testing, cybersecurity fundamentals, digital forensics, secure system administration, vulnerability assessment and risk management, cybersecurity laws and ethics, and incident response and threat analysis.

Table 6

Curriculum alignment and instructional design (Faculty Survey, Part I, n = 10)

Indicator	M	SD	Interpretation
1. Aligned with learning outcomes of the cybersecurity and network-security courses	5.90	0.32	Strongly Agree
2. Modules correspond to key curriculum topics	5.80	0.42	Strongly Agree
3. Guided exercises support competency-based objectives	5.80	0.42	Strongly Agree
4. Meaningful bridge between theory and practice	5.90	0.32	Strongly Agree
5. Design reflects industry-relevant ethical-hacking methodologies	5.90	0.32	Strongly Agree
6. Addresses the gap in hardware-based, experiential learning	5.80	0.42	Strongly Agree

RQ3. Development process. The development stages are described in Table 1. The expert panel (n = 13) rated the process at a composite of M = 5.91 (SD = 0.28; Table 7). Stage means ranged from 5.90 to 5.95. Four of the 15 indicators were rated 6.00 by all 13 respondents (SD = 0.00): the justification of hardware and software selection, the logical structure of the architecture, the alignment of scenario-based modules with ethical-hacking phases, and the traceable progression from specification to build. The lowest indicator was the adequacy of ethical and safety safeguards (M = 5.69, SD = 0.48).

Table 7

Expert validation of the development process (Validation Checklist, n = 13)

Stage	Stage M	Lowest-rated indicator (M)	SD = 0.00 items
Needs analysis	5.92	Requirements reflect intended outcomes (5.85)	1
Design	5.90	Ethical and safety safeguards (5.69)	2
Prototyping and development	5.95	Iterative revisions documented (5.92)	1
Testing and pilot evaluation	5.90	Testing adequate to identify issues (5.85)	0
Refinement and iteration	5.90	Development cycle documented (5.85)	0
Composite (15 indicators)	**5.91**	**SD = 0.28**	**4**

Note. *Stage means are the averages of the three item means in each stage. All stage means fall in "Strongly Agree."*

RQ4. Practical understanding and engagement. Before the pilot, readiness items were lowest for prior experience with cybersecurity tools (M = 3.86, SD = 1.77) and familiarity with basic ethical-hacking concepts (M = 3.98, SD = 1.58) and highest for motivation to learn beyond lectures (M = 5.05, SD = 0.96; Table 8). After the pilot, students (n = 38) rated all ten outcome items between 5.29 and 5.76 (composite M = 5.60, SD = 0.71). The highest were that hands-on activities connected theory to real-world scenarios (M = 5.76), that worksheets and debriefs reinforced learning (M = 5.74) and that more was retained from hands-on activities than from lectures alone (M = 5.71). The lowest and most dispersed item was improved confidence in performing cybersecurity tasks (M = 5.29, SD = 1.06). Faculty observation was similar (composite M = 5.72, SD = 0.45), highest for greater engagement than in traditional lectures and for responsible, ethical application of concepts (both 5.80). Baseline and post-pilot items are not identical and responses were anonymous, so the comparison between them is descriptive only.

Table 8

Baseline readiness, post-pilot outcomes and faculty observation (RQ4)

Indicator	M	SD	Interpretation
A. Baseline readiness (Pre-Implementation Survey, Part II, n = 42)			
1. Familiar with basic ethical-hacking concepts	3.98	1.58	Slightly Agree
2. Prior experience with cybersecurity tools or platforms	3.86	1.77	Slightly Agree
3. Confident performing basic network analysis	4.31	1.42	Slightly Agree
4. Understands ethical and legal boundaries of penetration testing	4.69	1.16	Agree
5. Prepared for hardware-based cybersecurity learning	4.26	1.48	Slightly Agree
6. Motivated to learn beyond lecture content	5.05	0.96	Agree
B. Post-pilot outcomes (Post-Implementation Student Survey, Part I, n = 38)			
1. Deepened practical understanding of concepts	5.61	0.55	Strongly Agree
2. Hands-on activities connected theory to real-world scenarios	5.76	0.49	Strongly Agree
3. Confidence performing cybersecurity tasks improved	5.29	1.06	Strongly Agree
4. Actively engaged throughout activities	5.66	0.71	Strongly Agree
5. Scenario-based modules effective in illustrating concepts	5.61	0.59	Strongly Agree
6. Better understood ethical and legal considerations	5.58	0.98	Strongly Agree
7. Retained more from hands-on activities than lectures alone	5.71	0.61	Strongly Agree
8. Increased interest in pursuing cybersecurity further	5.47	0.73	Strongly Agree
9. Worksheets and debriefs reinforced learning	5.74	0.45	Strongly Agree
10. Overall, improved practical competencies	5.55	0.65	Strongly Agree
C. Faculty observation (Faculty Survey, Part II, n = 10)			
1. Greater engagement than traditional lectures	5.80	0.42	Strongly Agree
2. Improved understanding of key concepts after exercises	5.70	0.48	Strongly Agree
3. Activities challenged students to think critically	5.70	0.48	Strongly Agree
4. Students applied concepts responsibly and ethically	5.80	0.42	Strongly Agree
5. Supported students at varying knowledge levels	5.60	0.52	Strongly Agree

Note. *Composites: A = 4.36 (SD = 1.47); B = 5.60 (SD = 0.71); C = 5.72 (SD = 0.45).*

RQ5. Usability, scalability and sustainability. Students rated usability and related indicators between 5.76 and 5.89 (composite M = 5.83, SD = 0.38), highest for the appropriateness of the challenge level (M = 5.89). Faculty rated scalability and sustainability items between 5.70 and 5.90 (composite M = 5.81, SD = 0.39), highest for cost feasibility and for the open-source, modular design (both 5.90) and lowest for ease of deployment (M = 5.70; Table 9).

Table 9

Usability, scalability and sustainability (RQ5)

Indicator	M	SD	Interpretation
A. Students (Post-Implementation Survey, Part II, n = 38)			
1. Easy to set up and operate during lab activities	5.82	0.39	Strongly Agree
2. Instructions and guided materials were clear	5.76	0.43	Strongly Agree
3. Interface and hardware appropriate for an academic setting	5.76	0.43	Strongly Agree
4. Scenarios appropriately challenging for skill level	5.89	0.31	Strongly Agree
5. Could be used effectively in other PUP courses	5.87	0.34	Strongly Agree
6. Could be replicated without expensive infrastructure	5.87	0.34	Strongly Agree
B. Faculty (Faculty Survey, Part III, n = 10)			
1. Easy to deploy and configure in a classroom or lab	5.70	0.48	Strongly Agree
2. Integrates into course schedules without disruption	5.80	0.42	Strongly Agree
3. Hardware and software accessible for SUC environments	5.80	0.42	Strongly Agree
4. Can be scaled across sections and courses	5.80	0.42	Strongly Agree
5. Potential for long-term institutional adoption	5.80	0.42	Strongly Agree
6. Cost reasonable and feasible for SUCs	5.90	0.32	Strongly Agree
7. Faculty with basic knowledge could facilitate without extensive training	5.80	0.42	Strongly Agree
8. Open-source, modular design supports future updates	5.90	0.32	Strongly Agree

Note. *Composites: A = 5.83 (SD = 0.38); B = 5.81 (SD = 0.39).*

Table 10

Feasibility of the proposed implementation plan (Feasibility Survey, n = 10)

Indicator	M	SD	Interpretation
Deployment readiness and resource planning			
1. Identifies hardware, software and facility requirements	5.80	0.42	Strongly Agree
2. Realistic timeline for course integration	5.80	0.42	Strongly Agree
3. Addresses faculty training and orientation	5.90	0.32	Strongly Agree
4. Feasible budget and cost allocation for an SUC	5.50	0.71	Strongly Agree
Institutional integration and scalability			
1. Outlines steps for scaling across sections and courses	5.80	0.42	Strongly Agree
2. Considers technical support to sustain operations	5.80	0.42	Strongly Agree
3. Phased pilot-to-full rollout is logical and achievable	5.80	0.42	Strongly Agree
4. Addresses technical and logistical risk mitigation	5.70	0.67	Strongly Agree
Monitoring, evaluation and sustainability			
1. Clear mechanisms for monitoring effectiveness	5.80	0.42	Strongly Agree
2. Supports long-term adoption beyond the pilot	5.80	0.42	Strongly Agree
3. Realistic given current resources and constraints	5.90	0.32	Strongly Agree
4. Overall, feasible for adoption at PUP	5.80	0.42	Strongly Agree
Composite	**5.78**	**0.45**	**Strongly Agree**

RQ6. Implementation-plan feasibility. Faculty and administrators (n = 10) rated the four-phase plan at a composite of M = 5.78 (SD = 0.45; Table 10). Budget and cost allocation was the lowest and most dispersed item (M = 5.50, SD = 0.71), while faculty training and orientation, and realism given current resources, were highest (both 5.90). In an open-ended comment, the chairperson identified five support needs: regular funding for updates and maintenance, faculty training, laboratory resources and technical support, integration into the cybersecurity curriculum, and continued collaboration with industry partners.

DISCUSSION

Principal findings. Across five instruments and four respondent groups, the toolkit was perceived as aligned with the curriculum, developed systematically, engaging, usable and feasible to sustain, with every post-pilot and expert composite between 5.60 and 5.91 on a six-point scale. Before the pilot, students' ratings pointed to a relative rather than an absolute shortfall: general hands-on activity was rated higher (M = 5.10) than access to hardware-based tools and laboratory resources (both M = 4.19), and students agreed (M = 5.02) that a gap exists between classroom knowledge and industry skills. This pattern is consistent with Philippine evidence of a gap between academic preparation and industry expectations (Miranda et al., 2025) and of institutional capability constraints in SUCs (De Ramos & Esponilla, 2022). It comes from one section, however, and cannot show how widespread the shortfall is at PUP. The large dispersion on the hardware items (SD = 1.55–1.70) indicates that students' experiences differed, and the contrast between the general and hardware-specific items may partly reflect item wording; brief cognitive interviews would clarify this.

Interpreting perceived gains. Students rated post-pilot practical understanding and engagement at 5.60, and faculty independently rated observed engagement at 5.72. The highest student ratings concerned linking theory to real-world scenarios and the value of worksheets and debriefs, which fits the reflective-observation and conceptualization stages of Kolb's (1984) cycle and the demonstration–practice–debrief structure of the sessions. Comparable studies point in the same direction: students who had never done a hands-on cyber-physical security lab rated one highly (Chi et al., 2022), and perceived confidence rose after a hands-on ESP32 project cycle (Zambrano-Mieles et al., 2026). These parallels concern direction rather than magnitude, since scales and designs differ. The present data do not demonstrate learning. Pre- and post-pilot items measure different constructs (baseline readiness versus perceived gains), responses were unmatched, and no test of change was computed. Higher post-pilot ratings may reflect real gains, but also novelty, instructor enthusiasm, social desirability (the researcher who developed the toolkit was involved in delivering the pilot) and ceiling effects, and this design cannot separate them. Improved confidence was the lowest and most dispersed outcome (M = 5.29, SD = 1.06), which suggests that benefits were not uniform; stratifying future analyses by prior experience would test this.

Usability, scalability and sustainability. Students and faculty rated these qualities at 5.83 and 5.81. Usability was measured with purpose-built items rather than a standardized instrument; Criollo-C et al. (2024) show how the CSUQ and NASA-TLX can add comparable usability and workload evidence, and they would be a logical addition. The scalability and long-term adoption ratings are prospective judgments, because the toolkit was used in one section and no other instructor, section or campus has yet used it. In Rogers's (2003) terms, the pilot demonstrates trialability and observability in a favorable setting, while compatibility and complexity for other instructors remain untested.

Implementation feasibility. Respondents rated the four-phase plan as feasible (M = 5.78), with budget and cost allocation the lowest and most dispersed item (M = 5.50, SD = 0.71). This is where theory and context both predict friction:

SUC acquisitions are governed by public-bidding rules (Republic Act No. 9184, 2003), and connectivity and awareness have been reported as adoption barriers (Alimboyong & Bucjan, 2021). The chairperson's emphasis on funding, training, laboratory support and curriculum integration reflects the facilitating conditions that Venkatesh et al. (2003) identify as a separate determinant of use. In Fixsen et al.'s (2005) terms, the pilot corresponds to initial implementation; installation (trained trainers, maintenance budget, procurement lead time) and full implementation remain proposals. Coburn's (2003) criteria of depth, sustainability, spread and shift in ownership offer a way to evaluate those later phases. Favorable ratings of a plan are not evidence that the plan will work.

Ethical and safety considerations. The lowest of the 15 expert-rated indicators was the adequacy of ethical and safety safeguards (M = 5.69, SD = 0.48), still "Strongly Agree" but below all other design indicators. Because the toolkit demonstrates techniques that are unlawful outside authorized settings, future versions should document safeguards explicitly, including authorization procedures, isolated test devices, supervision ratios and defensive debriefs. Mapping exercises to work roles and defensive tasks, as in the NICE Framework (Petersen et al., 2020), would help balance offensive demonstrations with mitigation.

Limitations. The pilot involved a class of 42 students enrolled in a Networks and Security course, with no comparison group or randomization. All outcome data are perceptions; no objective measure of skill was collected. The anonymous pre- and post-pilot surveys could not be matched, and 38 of 42 students completed the post-pilot survey because of absences. Ratings were uniformly high, so ceiling effects are possible. The toolkit's developer took part in delivering the pilot, which may have influenced responses. Reliability coefficients are not reported, and only descriptive statistics were computed. Exercises with 2.4 GHz radios required authorization, which may limit replication elsewhere. The findings are best read as early, setting-specific evidence, and the evaluations proposed in the Recommendation section would address most of these points.

Conclusion. This study developed a portable, ESP32-based ethical hacking toolkit and piloted it with select Computer Engineering students at PUP, Manila. Before the pilot, students rated hardware access and laboratory resources as the weakest part of their cybersecurity instruction. After three sessions, every group that evaluated the toolkit rated it very favorably. Faculty judged it well aligned with course outcomes (M = 5.85), IT and network-security professionals rated its development process as systematic and well documented (M = 5.91), and students and faculty reported higher engagement and practical understanding (M = 5.60 and 5.72). Students and faculty also rated it usable and scalable, and faculty and administrators judged the proposed implementation plan feasible (M = 5.78), with budget the main area to plan for. Together, these results show that a low-cost, portable toolkit can be built, delivered in class and well received in a resource-constrained setting. As a single-section pilot based on perceptions, the study establishes acceptability and feasibility. The next step is to measure learning directly, through objective assessments across more sections and institutions.

Recommendation. Future work should (1) add objective measures, such as a practical-task rubric and knowledge tests given before and after use, with matched identifiers; (2) use comparison groups or multiple sections and institutions; (3) report reliability coefficients and add standardized usability and workload instruments (Criollo-C et al., 2024); (4) publish a bill of materials and maintenance costs to substantiate affordability; (5) formalize the implementation plan by training additional faculty trainers, scheduling procurement ahead of the semester of use, and mapping modules to program outcomes (CHED, 2017) and NICE work roles (Petersen et al., 2020); and (6) extend the toolkit to defensive-security exercises document its safeguards in more detail.

DISCLOSURES

Author contributions. John Paul F. Avila: Conceptualization, Writing - Original Draft, Methodology, Funding, Data Curation, Writing - Review & Editing | Nelson C. Rodelas: Supervision, Writing - Review & Editing.

Conflict of interest. The authors declare no conflict of interest.

Funding source. This research received no external funding.

Artificial intelligence use. Authors used Claude AI for drafting and getting suggestions when editing the contents of paper. The authors reviewed and approved all content.

Ethics approval statement. This study was reviewed and cleared by the Research and Extension Office of the Polytechnic University of the Philippines Graduate School (Ethics Review Clearance No. 2026-377, issued 30 June 2026). Participation was voluntary, responses were anonymous, and data were used solely for academic and research purposes.

Data availability statement. The dataset will be available upon request from the researchers. data includes anonymous responses as requested by respondents of the study, thus information like name, age, gender, or any other identifying details won't be shared.

Acknowledgement. The authors thank the faculty, students, IT and network-security professionals and administrators of the Polytechnic University of the Philippines who took part in the study, and the student mentees who assisted during development.

Publisher's disclaimer. The views expressed in this article are those of the authors and do not necessarily reflect the views of the publisher. The publisher disclaims any responsibility for errors or omissions.

REFERENCES

Alimboyong, C. R., & Bucjan, M. E. (2021). Cloud computing adoption among state universities and colleges in the Philippines: Issues and challenges. *International Journal of Evaluation and Research in Education, 10*(4), 1455–1461. https://doi.org/10.11591/ijere.v10i4.21526

Chi, H., Liu, J., Xu, W., Peng, M., & DeGoicoechea, J. (2022). Design hands-on lab exercises for cyber-physical systems security education. *Journal of the Colloquium for Information Systems Security Education, 9*(1), Article 8. https://doi.org/10.53735/cisse.v9i1.140

Coburn, C. E. (2003). Rethinking scale: Moving beyond numbers to deep and lasting change. *Educational Researcher, 32*(6), 3–12. https://doi.org/10.3102/0013189X032006003

Commission on Higher Education. (2017). *Policies, standards, and guidelines for the Bachelor of Science in Computer Engineering (BSCpE) program* (CHED Memorandum Order No. 87, s. 2017). Republic of the Philippines.

Crawley, E. F., Malmqvist, J., Östlund, S., & Brodeur, D. R. (2007). *Rethinking engineering education: The CDIO approach*. Springer.

Criollo-C, S., Guerrero-Arias, A., Buenaño-Fernández, D., & Luján-Mora, S. (2024). Usability and workload evaluation of a cybersecurity educational game application: A case study. *IEEE Access, 12*, 12771–12784. https://doi.org/10.1109/ACCESS.2024.3352589

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology.

MIS Quarterly, 13(3), 319–340. https://doi.org/10.2307/249008

De Ramos, N. M., & Esponilla, F. D., II. (2022). Cybersecurity program for Philippine higher education institutions: A multiple-case study. *International Journal of Evaluation and Research in Education, 11*(3), 1198–1205. https://doi.org/10.11591/ijere.v11i3.22863

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature*. University of South Florida, National Implementation Research Network.

Karagiannis, S., Magkos, E., Karavaras, E., Karnavas, A., Nikiforos, M. N., & Ntantogian, C. (2024). Towards NICE-by-design cybersecurity learning environments: A cyber range for SOC teams. *Journal of Network and Systems Management, 32*(2), Article 42. https://doi.org/10.1007/s10922-024-09816-w

Katsantonis, M. N., Manikas, A., Mavridis, I., & Gritzalis, D. (2023). Cyber range design framework for cyber security education and training. *International Journal of Information Security, 22*(4), 1005–1027. https://doi.org/10.1007/s10207-023-00680-4

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Prentice-Hall.

McKenney, S., & Reeves, T. C. (2019). *Conducting educational design research* (2nd ed.). Routledge.

Miranda, J. P. P., Tayag, M. I., & Canlas, J. D. (2025). Cybersecurity skills in new graduates: A Philippine perspective. *International Journal of Advances in Applied Sciences, 14*(4), 1217–1228. https://doi.org/10.11591/ijaas.v14.i4.pp1217-1228

Peffers, K., Tuunanen, T., Rothenberger, M. A., & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of Management Information Systems, 24*(3), 45–77. https://doi.org/10.2753/MIS0742-1222240302

Petersen, R., Santos, D., Smith, M. C., Wetzel, K. A., & Witte, G. (2020). *Workforce framework for cybersecurity (NICE Framework)* (NIST Special Publication 800-181 Rev. 1). National Institute of Standards and Technology. https://doi.org/10.6028/NIST.SP.800-181r1

Republic Act No. 9184. (2003). *An act providing for the modernization, standardization and regulation of the procurement activities of the government and for other purposes* (Government Procurement Reform Act). Republic of the Philippines.

Rogers, E. M. (2003). *Diffusion of innovations* (5th ed.). Free Press.

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Vykopal, J., Čeleda, P., Šeda, P., Švábenský, V., & Tovarňák, D. (2021). Scalable learning environments for teaching cybersecurity hands-on. In *2021 IEEE Frontiers in Education Conference (FIE)* (pp. 1–9). IEEE. https://doi.org/10.1109/FIE49875.2021.9637180

Zambrano-Mieles, J., Tupac-Yupanqui, M., Mari-Loardo, S., & Vidal-Silva, C. (2026). Integrating ESP32-based IoT

architectures and cloud visualization to foster data literacy in early engineering education. *Computers, 15*(1), Article 51. https://doi.org/10.3390/computers15010051